DOCUMENT RESUME

ED 180 478 IR 007 952

AUTHOR	Hitchingham, Eileen E.
TITLE	A Study of the Relationship between the Search Interview of the Intermediary Searcher and the Online System User, and the Assessment of Search Results as Judged by the User. Final Report.
INSTITUTION	Oakland Univ., Rochester, Mich. Kresge Library.
SPONS AGENCY	Bureau of School Systems (DHEW/OE), Washington, D.C. Office of Libraries and Learning Resources.
BUREAU NO	475AH70111
PUB DATE	Aug 79
GRANT	G007702309
NOTE	107p.
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	*Information Retrieval; On Line Systems; *Question Answer Interviews; Questioning Techniques; *Relationship; *Relevance (Information Retrieval); Search Strategies; Surveys; Tables (Data); *User Satisfaction (Information)

ABSTRACT

Using data gathered from 54 evaluation forms and 18 recorded interviews, this study focused on in-person users of the MEDLINE system at three search sites to determine the relationship between presearch interviews and user assessments of precision and concern for recall in online searches, and to determine differences in assessment characteristics exhibited by student and faculty users of the system. Relationships were demonstrated between the value designation and the user's satisfaction with the proportion of relevant citations, the user's concern for recall, the relevance score, and the perception of the searcher in the interview. User information-giving during the interviews was related to relevance scores for the searches, and question-asking by the searcher was related to information-giving by the user. Faculty and student users showed no significant differences in their assessments of value, relevances, satisfaction with the proportion of relevant citations retrieved, and in their perceptions of the searcher in the interview. Appendices to the study include samples of the tools used and a bibliography. (FM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED180478

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Final Report

Project No. 475 AH 70111
Grant No. G 00 7702309

A STUDY OF THE RELATIONSHIP BETWEEN THE SEARCH INTERVIEW OF THE INTERMEDIARY SEARCHER AND THE ONLINE SYSTEM USER, AND THE ASSESSMENT OF SEARCH RESULTS AS JUDGED BY THE USER

Eileen E. Hitchingham

Kresge Library
Oakland University
Rochester, Michigan 48063

August 1979

The activity which is the subject of this report was supported in whole or in part by the U.S. Office of Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred.

U.S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE

Office of Education
Office of Libraries and Learning Resources

IR007952

Preface

Many people have been of assistance in this study. My most particular thanks are extended to the participating searchers at the Shiffman Medical Library of Wayne State University, the Medical Center Library of the University of Michigan, and the Jesse H. Jones Library of the Houston Academy of Medicine - Texas Medical Center Library. Their suggestions in the preliminary stages were most valuable; their cooperation in incorporating study procedures into their already busy schedules is most appreciated. Special thanks are due to the directors of these libraries for fostering an atmosphere which encourages cooperation in research efforts such as this one. For her immediate interest in the project and her assistance in making long-distance data collection run smoothly, I am particularly grateful to Sara J. Jackson, Assistant Director for Public Services, at the Jesse H. Jones Library.

Mary Kitto of Oakland University has been involved in this study since its inception. I am grateful not only for her technical assistance in setting up forms, arranging meetings and typing this report, but also for the very special interest and support she has provided throughout the study.

The work of Susan Woodford and Mary Pensyl of the MIT Science Library in coding the interviews was most valuable. Their expert knowledge of the data base and the interview situation facilitated the process throughout. Margaret K. Park of the Office of Information Systems, University of Georgia was most helpful in providing copies of transcripts for practice coding sessions.

TABLE OF CONTENTS

Page

Preface i

List of Figures iv

List of Tables v

I. Introduction 1

- Purpose 1
- Background 1
- Problem Rationale 2

II. Review of the Literature 6

- Assessments and Users 6
- The Search Interview 16

III. Methods and Procedures 22

- Design 22
- Online Data Base 22
- Sampling Plan 23
- Data Collection Instruments 26
- Data Collection 30
- Data Treatment 33

IV. Results 35

- Data Sources 35
- Overall Characteristics 35
- Search Results 37
  - Interrelationships 38
  - User Differences 47
  - Information-Giving, Question-Asking 53
  - Interview Profiles 54

V. Conclusions 61

Bibliography 63

Appendices

A. Consent Form 67
B. Request Form 70
C. Evaluation Form 78
D. Procedures 83
E. Cover Letters 87
F. Coding Reliability 89
G. User Requests: Analyzed Interviews 92
H. Coded Interview 96

LIST OF FIGURES

Page

1. Categories in the Interaction Process Analysis Scheme 29

2. Value of MEDLINE Search Results in Meeting the Need Prompting the Search Request 39

3. Distribution of Relevance Scores for MEDLINE Searches 41

4. Distribution of Relevant Citations in the MEDLINE Searches 42

5. Distribution of Mean Scores for User Perceptions of the Search Interview 44

6. Distribution of Concern for Recall Scores 45

7. Distribution of Satisfaction with the Proportion of Relevant Scores 46

8. Percent of Total Interaction by Category for 18 Transcribed MEDLINE Interviews 55

9. Searcher and User Interaction by Categories as Percent of Total Interaction for 18 Transcribed MEDLINE Interviews 57

10. Searcher Interaction by Category as Percent of Total Searcher Interaction by Type of Interview 58

11. User Interaction by Category as Percent of Total User Interaction by Type of Interview 59

LIST OF TABLES

Page

1. Correlation of Evaluative Measures with User Designations of Value 38

2. Summary Results for Testing Differences by Type of User (Faculty, Student) 47

3. User Type by Regular Reading of Journals 49

4. User Type by Publication of Related Article(s) 49

5. Summary Results for Testing Differences by Subject Area of Search (Biomedical, Social Science) 50

6. Summary Results for Difference by Previous Online Interview Experience 51

7. Summary Results for Testing Differences by Type of Interview Session (Terminal, Non-Terminal) 52

8. Correlation of User Information-Giving with Relevance, Concern for Recall, and Searcher Question-Asking 53

I. Introduction

Purpose

The problems investigated in this study are specified within the framework of four questions:

1. When a user of an online system reviews the results of the search, is the user's designation of the value of the search (the ability of the results to meet the need prompting the search request) related to the user's designation of relevance, satisfaction with the proportion of relevant citations, concern for recall, and the perception of the searcher in the search interview?
2. Do specified user groups (students and faculty) differ in their assessment of search parameters and in their participation in the search interview?
3. In the search interview, is information-giving activity on the part of the user related to the user assessment of results in the areas of relevance and concern for recall?
4. If user information-giving activity is related to user assessments of relevance and concern for recall, is this relationship enhanced by question-asking activity by the searcher?

Supplementary factors related to these questions are also considered.

Background

Growth in the accessing of online bibliographic data bases appears to be a continuing phenomenon. Over a period of four years the number of such searches almost tripled, from 700,000 in 1974 to 2,000,000 in 1977 (Williams, 1977). Williams considered several factors to be significant in supporting this growth, (1) increased use of the National Library of Medicine (NLM) System; (2) the emergence, in 1977, of a new online vendor, Bibliographic Retrieval Services (BRS); (3) increased exposure of students to online system in higher education; and (4) decreasing costs per record searched.

Interest and activity related to the provision of online search services is high. The Reference and Adult Services Division of the American Library Association has formed a special discussion group on machine assisted reference services ("Librarians on MARS," 1974).

A conference on The Online Revolution in Libraries, scheduled by the University of Pittsburgh, and for which 300 attendees were anticipated, drew more than 700 interested librarians and library educators (Nyren, 1978). The advent, in 1977, of two new journals devoted to online access to bibliographic data bases, ONLINE and Online Review, provides support for the belief that the growth of online users will be sustained.

Since it is likely that online services will become more common in libraries and information centers in the future, it appeared appropriate to examine aspects related to the effectiveness of these services in meeting information needs.

Within the confines of this study effectiveness measures are user designations in the areas of search value, relevance, satisfaction with the proportion of relevant citations, concern for recall and perception of the searcher in the search interview. Aspects specifically examined are interrelationships of user assessment designations, differentiation in assessment and search interaction by types of users (students and faculty), and an examination of theoretically significant roles of the user and the searcher in the search interview and their subsequent relationship to effectiveness measures. Excluded from this study is any specific examination of actual search strategies performed for an individual search, although this obviously has aconsiderable impact on search results. However, the focus outlined here (user assessments and interview factors) appears to contribute, in part to supplying information for gaps suggested by Swanson (1975) which impede a better understanding of the complex search process.

> Studies of personal, largely verbally supplied services such as answers to reference queries and search strategy assistance are scarce.... Could guidelines be provided to information personnel to maximize their performance in this communication process? The question cannot be answered without appreciable data on characteristics of these communications including behavioral data for both patrons and information personnel. (p.148)

Problem Rationale

The questions outlined as the purpose of this study suggest that three areas are of interest, (1) the interrelationship of assessment measures, (2) assessment differences exhibited by different types of users, and (3) the search interview.

Interrelationships. Even though varying question formats are used, a number of evaluation questionnaires used by institutions providing retrieval services appear to elicit evaluative information from the search user in both specific areas (e.g., relevance of citations, absence of known relevant citations, timeliness, feelings about the interview, amount of time saved), and in a more general overall area (e.g., value of the search, satisfaction with the search, helpfulness of the search). Four question types commonly occuring in these questionnaires are: of the citations retrieved how many did the user find to be relevant; did the user feel that the search missed any relevant citations; what did the user feel about the search interview; and did the user find the search results to be valuable (Daniels, 1978; Hitchingham, Note 1).

If overall evaluative assessments are considered to be represented by questions on previous evaluations which ask the user to respond, on a four- or three-point scale, whether the search was, for example, of Major Value, Considerable Value, Minor Value, No Value (Lancaster, 1972; Mick, 1977); or Very Useful, Of Some Use, Of Little Use, Of No Use (Carmon, 1975); or Very Helpful, Helpful, Moderately Helpful, Not Helpful (Tagliacozzo, 1977); or Very Satisfactory, Generally Satisfactory, Not Satisfactory (Benenfeld, Marcus, Pensyl, and Reintjes, 1975); it is noted that the majority of users (60-91%) respond by checking the first two most favorable categories. This suggests that users of retrieval systems will generally find that the results are valuable or satisfactory, whatever other assessments they make. From an interpretive and diagnostic viewpoint it is desirable to know how more specific user assessments relate to the overall assessment.

User Characteristics. Some users of online information retrieval systems approach the system with a greater background knowledge of the topic to be investigated in the search. This background knowledge is the result of a continuing involvement with the specific area of investigation, i.e., the search is initiated to supplement already existing knowledge of literature in the subject area. In his review of experimental relevance studies, Saracevic (1970, p. 137) concludes that greater subject knowledge on the part of the assessor leads to more stringent assessments of relevance. In the same vein, it might be expected that greater knowledge of the literature of a subject allows more stringent assessments of recall, i.e., the knowledgeable user can note, and is concerned with, a lack of completeness in search results. Knowledgeable users may also place more stringent demands on the searcher in regard to their assessment of the search interview.

Since, as previously noted, users of information systems tend to react favorably to questions concerning the overall value of the search, it was not believed that knowledgeability of the subject would provide a distinguishing factor when the overall assessment was considered. In a similar manner, it was not expected that knowledgeability would provide differentiation in the satisfaction with the proportion of relevant citations. This satisfaction level is more likely to indicate a tolerance to sifting the wheat from the chaff in the retrieval results, and thus it would appear to be more related to individual temperament than to knowledgeability.

In this study faculty members who used an online service for research purposes, where research was considered to relate to grant activity or the publication process, were considered to be literature knowledgeable users. Faculty members can specialize in a limited area and thus can be more familiar with that area. It is generally assumed that there is some pressure, if only in the tenure process, for faculty members to publish. Such pressure constitutes a job-related need on the part of faculty to be literature conscious. For example, in a study of medical school faculty publication rates (Pearse, Flora, Freeman, & Peeples, 1976), the authors note a peak in publications per person, per year, in the early 40's. They attribute this peak to variations on thesis work followed by a decline until a new research area is selected.

In contrast, most students, because of their role as students, have to focus on a number of subject areas. Subject areas may be related, but there is less opportunity, because of conflicting information needs, to concentrate in depth on specific topics, and to have a cumulative record of involvement with one area. Thus students, in general, were considered to be less literature knowledgeable.

The Interview. Among several themes noted in publications concerning the library reference interview process (and by correlation the interview process for online searching), an important role of the intermediary (searcher) as a question asker emerges. For example, Francillon (1959) indicates that "... the first question often does not express the real intent of the requester. It is often necessary for the librarian to ask other questions"(p. 193). Taylor (1968) states, "Reference librarians and information specialists have developed, both consciously and unconsciously, rather sophisticated methods of interrogating users" (p. 179). In an instructional module for negotiating the reference query, Jahoda (1975, p. 12) indicates that the reference librarian should use open questions in the initial negotiation stages, and employ closed questions at the final stage of negotiation.

It follows that if question-asking is an important role for the searcher, then information-giving actions are significant in the user's role in the interaction process. Tessier, Atherton, and Crouch (1977) suggest that during the interaction for planning a computer search "a user will provide an immense amount of information about his requirements, expectations, and compromises" (p. 386).

It was considered that the amount of information given by the user might impact on search results. Since this information could be either voluntary or searcher elicited, it was desirable to assess the relationship between the searcher's asking of questions and the user's giving of information, and the user's ultimate assessment of the results.

II. Review of the Literature

Sources relevant to the three areas considered in this study (user assessments, user differences, the interview) are noted in this section. Assessment parameters and users are treated together since studies intermix these aspects. The interview is considered in a separate section.

Assessments and Users

Although the relevance study conducted by Rees and Schultz (1967) was not an examination of an operational system, but rather an assessment of experimental relevance judgments (i.e., defined groups of individuals making relevance judgments on test documents) it is included in this review because of some suggestive results concerning differences in relevance assessments by groups of people making a relevance decision. As the authors note, "It was assumed that differences in the extent of subject expertise and experience in research would result in variations in conceptualization of the information need as revealed by differences in the relevance ratings" (Vol. 1, p.29).

Subjects were 184 judges divided into five expertise-related groups. The groups were comprised of (a) 40 medical experts in the field of diabetes (i.e., MD's, 14 of whom were researchers, and 26 who were involved in patient care or medical education); (b) 29 medical scientists (individuals with a Ph.D. and working in biomedical research); (c) 25 residents (MD's involved in clinical care of patients); (d) 29 second year medical students, assumed to have limited expertise and practice in biomedical research, and (e) 61 medical librarians with little, or no direct experience in biomedical research (p. 38). The judges evaluated 16 documents for their relevance to a detailed description of a research project concerning sugar transport in the intestine. The dependent, evaluative measures consisted of four questions directly considering relevance (overall relevance, relevance with respect to the formulation of the research problem, relevance with respect to the interpretation of the finds), and one question concerning the overall usefulness of the document with respect to the research.

Although the total study covered a number of aspects of the relevance judgment (e.g., relevance judgments in regard to different kinds of document representation — citation, abstract, full text; and relevance judgments made in response to hypothesized stages of the research problem — initial formulations, carrying out the experiment, and analysis of results), only those findings summarized in the conclusion which are related to the current study are considered here.

There were significant differences in the mean ratings assigned by judgment groups to the document set. Scientifically oriented groups (e.g., medical experts and medical scientists) assigned lower mean relevance ratings than did less scientifically oriented groups (e.g., residents and medical librarians) when the overall relevance assessments are considered (p. 271). A more refined breakout, analyzed earlier in the study, indicates that the actual ranking by group, of mean relevance ratings includes, from high to low, medical librarians, medical experts (non-researchers), residents, medical experts involved in research, and medical scientists (p. 179). Thus, mean relevance ratings appear to become more stringent as one continues along a scale from clinical to more direct research involvement with the problem.

A second finding of interest to the present study was the indication that, at least on a document to document basis, responses concerning the usefulness of the document were highly correlated to the responses to the question concerning overall relevance (p. 273). The aspect of this finding that is considered pertinent to the present study is whether a similar high correlation between the relevance figure determined for a search, and the user's determination of the value of the search in meeting the need prompting the request, would occur.

The AIM-TWX study conducted by Lancaster (1972) appears to be a seminal study in the evlauation of online searching because it is one of the first to explicate the idea that one of the best indications of the success of a search is the user's subjective assessment of the search's value. In addition, the four-point value scale (major value, considerable value, minor value, no value) developed by Lancaster appears to have influenced the decision of subsequent investigators (e.g., Carmon, 1975; Tagliacozzo, 1977; Mick, 1977; Jahoda, Bayer, and Needham, 1978) to adopt the same scale or to adapt the scale in a similar manner for their own needs.

In the AIM-TWX study 48 users conducted their own searches on the Abridged Index Medicus data base which was available for online searching at that time. Users were equally divided between those who had never used the system before and those who had used it at least once. Precision values were determined for each search by dividing the number of unique relevant citations noted by the user in the search results by the total number of unique citations printed (precision scores equate to the relevance score determined in the present study). The average precision figure for 45 searches was 63.1%. As Lancaster notes, this figure is higher than the average precision (50%) achieved in 299 off-line searches studied in his earlier MEDLARS investigation (1968). In the MEDLARS study precision figures are based on user evaluations of the actual documents (not just the citations); in the AIM-TWX study users were given copies of the actual articles if this was necessary for unequivocal relevance assessments.

Approximately one-third of the AIM-TWX users considered in the study were M.D.'s (associate and assistant professors, residents, interns and postdoctoral fellow); ten users were third or fourth year medical students. The remaining users included research associates and assistants, a physical therapist, and an executive from a company manufacturing medical equipment. More than half of the searches appear to have been initiated for a research paper, review article or book chapter; writing a book or thesis; or preparing a grant application).

Perhaps the most interesting aspect of the AIM-TWX study is a consideration of the value that users attached to the search results. Like users in a number of studies following AIM-TWX, the group as a whole is perceived as attaching a high value to their search output, i.e., for the 38 searches evaluated on the four-point scale, 67% were designated as being of major or considerable value.

Moll (1972) reports on the evaluative responses of 62 users reviewing 100 AIM-TWX or MEDLINE searches done at the University of Virginia from August 1971 through January 1972 (MEDLINE searching, providing greater access to the biomedical literature, replaced AIM-TWX searching of a more limited online data base during the period studied). More than 300 searches were conducted during the study period, thus the return rate for questionnaires appears to have been approximately 30%. For 82 of the evaluated searches, users indicated that the citations furnished were either "most useful," or of "considerable interst." For 18 of the evaluated searches the citations were considered to be of "little interest" or "worthless." It is noted that a total of 2,082 citations were produced for the 100 evaluated searches. For 55 of the searches an estimated 50% or more of the citations sent were considered by users to be helpful in their work. It is interesting to note that 27 of the searches for which users estimated that 25% or fewer of the citations supplied were helpful were still considered to be searches that were "most useful" or of "considerable interest." This suggests that, for some users, a low relevance score does not reflect upon the ultimate judgment of an overall value of the search.

In 1974, almost two years after MEDLINE services were initiated at the Calder Memorial Library of the University of Miami Medical School, users were surveyed concerning their evaluation of the service (McCarthy, Maccabee & Feng, 1974). Questionnaires were sent to 350 locatable previous users. These users had initiated more than 1,200 searches in the time-frame considered. Almost 50% of the questionnaires were returned.

While users were not asked to respond to a direct question concerning search value, the authors assume general user satisfaction from the responses noted for three questions. First, even though 73%

of the users indicated that their search had been paid for by institutional funding, 44% of the users indicated that they would use the service with the same frequency if they had to pay with their own funds. Second, 62% of the respondents said that they had recieved about the number of citations they wanted. Finally, users were asked to supply specific criticisms of the MEDLINE service; for almost 60% of the returned questionnaires no specific criticisms were noted.

The authors suggest that there may have been a difference in satisfaction related to the type of interview conducted (phone or in person), but other than one comment in this area from a user, they provide no data to support this speculation.

Benenfeld, Marcus, Pensyl and Raintjes (1975, pp. 6-1 to 6-3) report evaluations by users of online services at M.I.T. Under the auspices of the Northeast Academic Science Information Center (NASIC), an experimental, pilot operation of a computer-based reference search service was initiated at the M.I.T. libraries in November 1973. The user population included academic users (faculty, graduate and undergraduate students, staff), and industrial users. In December 1974 an evaluation questionnaire was sent to all users of the service from the time of its initiation. One questionnaire was sent to each user, some users receiving the questionnaire may have had more than one search request. Almost 50% of the questionnaires were returned.

Through NASIC, M.I.T. users had access to a range of data bases in science and technology, the social sciences and the humanities. MEDLINE services were available through the National Library of Medicine. Approximately one-third of the more than 300 searches conducted for the period March 1974 through February 1975 were MEDLINE searches (p. G-1).

Responding users were, in general, considered to be satisfied with the service. More than 90% of the users responding on a three-point satisfaction scale (very satisfactory, generally satisfactory, not satisfactory) indicated that they were either generally satisfied (41%) or very satisfied (50%). In response to a four-point-scale question concerning citation relevance (high relevance, moderate relevance, marginal relevance, no significant relevance) most users of the service indicated moderate (39%) to high relevance (45%) for the citations they received when these citations were considered in relation to their initial problem.

Almost 35% of those responding did not choose to designate their status (e.g., faculty or student), thus it is not possible to examine user responses in relation to status for this study.

Jestes (Note 2, 1974) reports on evaluations of an online search service directed to the CAIN (Bibliography of Agriculture) data base.

In 1973 a grant to support an experimental service to provide citations from CAIN was awarded to the Library of the University of California at Davis. The report notes implementation procedures and costs as well as user evaluations collected after three months of service operation.

The population served included researchers in departments and laboratories at the campus, and Agricultural Extension Specialists who attended a demonstration at the university. The service was free to users, but they were given an indication of the search cost when they received their results. Questionnaires were sent with each bibliography; of a total of 237 questionnaires sent, 87 (37%) were returned.

Responses for three of the questions are considered here. Relevance of citations is reported as a total for all citations received by respondents completing the questionnaire. Of a total of 10,333 citations sent, 36% were considered to be directly pertinent to the research problems of the user, 25% were thought to be interesting but not directly pertinent, and 37% were considered not useful (2% of the citations were not accounted for). Even though the percent of directly relevant citations was not particularly high for this study, users gave some indications that they were satisfied with the results. Almost 90% of the respondents indicated that the search was worth the cost ($11.61 was the average cost per bibliography), and 67% indicated that in the future they would be willing to pay for such a service.

This report is interesting because of the results concerning relevance and perceived satisfaction, i.e., a group comprised, for the most part, of academic researchers indicated a relatively low overall precision score, yet appeared satisfied with the results, so insofar as satisfaction can be equated with a willingness to pay for future services.

Hoover (1976) reports survey results from users of online services in another academic environment, the University of Utah libraries. In April-May 1975 all users received an evaluation questionnaire with their results. Of the 76 questionnaires distributed 26 were returned. The survey period occurred almost 15 months after the initiation of services at Utah.

Most respondents were faculty or graduate students. Sixty-four percent of the respondents indicated that they were first-time users. Multiple data bases were searched for a number of the requests. Users were asked to give an estimate of the percent of relevant citations in their results by checking a five-point-scale (e.g., 0-20%, 20-40%, . . . 80-100%). Fifty-four percent of those responding indicated that 60% or more of the citations were relevant. All of the respondents answered yes when queried whether the service was worthwhile. Hoover indicates that faculty groups had the highest success (i.e., more faculty members indicated 60% or greater relevance), and suggests that this occurred because they knew most precisely the topics they wanted to search. However, examination of the spread of responses shows that each user group, other than professional researcher, had low and high relevance searches of an almost equal number.

Faculty and student users of MEDLINE were surveyed in a university that did not have a medical school (Hitchingham, 1976). The study covered a period of four months. Twenty-one faculty and students made use of the service, which was provided free during that period. Thirty-six searches in five subject fields were initiated. Thirty-one of the searches were for a research related purpose (ongoing research, writing a research paper or review article).

Evaluation forms were sent with each of the completed searches. Twenty-seven (75%) of the evaluation forms were returned. All of the students returned forms, while only 13 out of 22 faculty searches were evaluated. Of the 27 searches evaluated 25 were found to be of major or considerable value. Users were asked to mark relevant citations on the printout, a precision value for each search was computed after examining the marked printouts. The average precision value for all searches was 53.9%. Precision values achieved for the searches had no apparent relationship to the amount that the user indicated she/he would pay for such a service in the future (i.e., low and high precision values were noted in conjunction with willingness to pay from $2.50 to $50 for future searches). Higher mean precision values were noted for student searches (60%) than for faculty searches (47.3%) and it was suggested that this might be indicative of less critical judgment on the part of students as to what is, or is not relevant to their topic.

Tagliacozzo (1977), in a follow-up study of MEDLINE users who had requested a search from one of seven midwestern centers during the period April-September 1973, achieved one of the highest response rates noted in reviewing online evaluations. Almost 90% of the 1,017 questionnaires reaching the addresses were returned. From an earlier study detailing the characteristics of all the MEDLINE users (Tagliacozzo, 1975, p. 295), these responses can be assumed to have come, for the most part, from in-person users of MEDLINE as opposed to requests received by phone, teletype or mail. It also appears that responses reflect both users who had one search, and users with more than one search (i.e., one user may have returned several evaluations). Almost half of the 810 search initiations were first-time users of MEDLINE.

Two questions directed to users asked that they consider the "helpfulness" of the MEDLINE search and the "usefulness" of search results. For the helpfulness aspect users were asked to indicate whether the search was not helpful, moderately helpful, helpful, or very helpful. For 895 searches evaluated in this item, 60% were considered to be helpful.

The second question asked the user to make a point on a line to indicate the usefulness of the search results. The line was anchored at the left by the words "completely useless" and at the right by "very useful." Later the line was divided into seven equal intervals to obtain scores (-3 to +3). Tagliacozzo suggests that this question required more precise judgment on the user's part than the "helpful" question. For the 826 searches evaluated on this question 69% were scored for points corresponding to +1 through +3.

When the responses (N = 821) to the two questions were compared by means of a frequency matrix of helpfulness responses versus line scores, the author notes that there were some ambiguities in the moderately helpful and helpful categories. For example, six responses indicate that the MEDLINE search was helpful, but a -1 to -3 score is noted for these same searches on the usefulness dimension. Response patterns noted for "moderately helpful" show an amost normal distribution when plotted against usefulness scores. Tagliacozzo suggests two interpretations of this occurrence, either the word moderately implied positive connotations to some users, and negative to others; or that users marking "moderately helpful" were "those who did not have -- or were unwilling to express -- a strong opinion on the MEDLINE service" (p. 246).

In addition to the helpfulness-usefulness comparison, the author contrasts the users' indications of the number of useful references retrieved with values noted on the helpfulness dimension (N = 706). Dichotomizing the helpfulness responses (higher = very helpful, helpful, lower = moderately helpful, not helpful) and the number of useful references indicated by the user (0-5, and 6 or more), she notes a strong association between the number of useful references retrieved and the judgment of helpfulness. More than 80% of the users with the higher numbers of relevant references also indicate that the search was helpful or very helpful; only 40% of those retrieving five or fewer references gave such indications.

An assessment aspect that might be considered to relate to a concern for recall by the user was also studied. For this contrast (N = 795) frequencies of dichotomized helpfulness values were related to the user responses indicating either that no relevant ciations were missed, or that one or more relevant citations were known to be missing. She notes a significant association in these measures, i.e., 70% of the responses indicating no missed citations were aligned with upper values on the helpfulness scale; only 49% of the users noting one or more missed citations chose the higher categories.

The author notes that users may give contradictory and dissonant responses when considering several aspects of an evaluative questionnaire. Furthermore, she suggests caution in interpreting from single responses, particularly those which elicit general or overall judgments, that the user's information needs were, or were not satisfied by the service.

Relatively few evaluation studies have been reported for users in a non-academic environment. Fosdick (1977) reports the responses of online users at the U.S. Army Construction Engineering Laboratory (CERL) in Champaign, Illinois. Evaluation forms were sent to 27 users noted as recipients of online searches over a two-year period; 26 of the forms were returned. Fosdick used the same scale for relevance as that previously noted in Hoover's study (1976). Unlike the academic users, respondents at CERL indicated generally lower estimates of citation relevance. Almost 70% of the users indicated that 40% or fewer of the citations were relevant. However, in response to a broader question concerning the usefulness of the service, CERL users evidenced the same favorable trend apparent in many other online evaluations. Twenty-five of the respondents indicated that they felt that the service was a useful addition to other library services; twenty users believed that the search was worth the cost.

A second study focusing on non-academic online searching is the evaluative component of the final report on the Lockheed Public Library Experiment (Mick, 1977). The project involved development and implementation of a program to assess the viability of providing access to major online bibliographic data bases in the public library setting. Only those evaluative aspects related to the current study are considered here.

During the first year of the project services were provided free to users of four California public libraries. In the second year one-half of the actual costs (terminal connnect time and printing costs) were charged back to the user. Two libraries continued with the project in the third year. During this period the fee system was based upon total recovery of connect and printing costs.

In the first year there were 1,236 patrons who requested a search. For the second year (recovery of half the costs) usage dropped; in this period there were 611 users. In the third year (full cost recovery, only two libraries participating) 326 users requested an online search. Students (graduate and undergraduate), educators (teachers, professors, school administrators) and scientists/research personnel accounted for the bulk of users over the three-year project. Use by graduate students and educators increased over the time span encompassed by the project, while use by undergraduate students and scientists declined as half, and full cost fees were introduced.

Although the details of eliciting evaluative information from users are not explained, it appears that a questionnaire was sent to users at some time during the study. Response rates appear to have been fairly low, i.e., evaluative responses noted respresent 36% of the first year users, 26% of the second year users, and 29% of the third year users. Responses to questions concerning the search value, the adequacy of results, and the number of citations considered useful are discussed here.

In general, responding users showed little change over the three-year period in their perception of the value of the search. When asked to make a value designation on a four point scale (major value, considerable value, minor value, or no value), 68% to 76% of the users over the three years indicated that the search was of major or considerable value.

In response to a yes/no choice concerning whether the results adequately answered the question, 53% to 56% of the users responded in a positive mode. Relevance determinations for citations were not given, however, in reviewing the results of Mick's study it is possible to speculate that many users were satisfied with a fairly low proportion of useful articles (e.g., almost 75% of the third year users found 0 to 15 citations to be useful). An earlier, interim report of the public library experiment (Summit & Firschein, 1976, pp. 3-3) indicates that in the second year of the study an average of 88 citations per search were printed offline. Mick reports that 69% of the second year users had 0 to 15 useful citations.

The results from this study suggest that whatever the other factors involved, users will give rather favorable responses to an overall question concerning the value of the search.

One of the few situations in which online users in two different types of organizations were studied has been reported by Jahoda, Bayer, and Needham (1978). The results presented were obtained from records maintained for each search, and user feedback obtained for most of the searches. Participants in the study included 50 individuals from the Chemistry Department at Florida State University (faculty, graduate students at the dissertation stage, and research associates), and 234 scientists and technologists at the Monsanto Textile Company. Services were free to users during the 13 month period reported. Results from 353 FSU searches and 345 Monsanto searches are reported for the free period. Findings related to the present study are examined here.

Results from the two groups were contrasted. Academic users were found to be more interested in an exhaustive approach to the literature (everything available) than were the industrial users. Monsanto users were more likely to approach the service because they were looking for specific facts and procedures, When the user analysis of results is considered, the authors note that both FSU and Monsanto users were generally pleased. Users were asked whether the number of citations they received was "about right," "too many" or "not enough." For 68% of the FSU searches and 64% of the Monsanto searches, users indicated that they had received about the right number. The average or median number of total citations retrieved for the searches is not reported, although it is known that for the FSU searches only 25% resulted in more than 21 online citations being printed, and there were offline prints for 36% of the searches. Sixty-three percent of the Monsanto searches had offline citations. It seems that even though FSU users are characterized as being more exhaustive in their approach to the literature, they may have received search results which, on the whole, included fewer citations than those received by Monsanto users. There was little concern by either group that too many citations had been retrieved (7% of the searches at FSU, 5% at Monsanto); some concern that not enough citations were presented was more likely to be expressed (13% at FSU, 19% at Monsanto).

There were no significant differences in the two groups when they reported their perception of the overall utility of the searches. The theme of generally favorable responses to an overall evaluative question noted previously in this review, is repeated for both user groups. Users were asked to indicate whether the searches were "very useful," "of some use," "of marginal use," or "of no use." FSU users found that 78% of the searches were either very useful or of some use, at Monsanto 75% of the searches were in these categories.

Student users of online services were the focus of interest in a study at the University of Delaware (Kobelski & Trumbore, 1978). The groups studied were satisfied and non-satisfied student users who had received a cost-subsidized search during the 1976-77 academic year. The period covered was the second year of operation for the search program. The subsidization system meant that students were charged half the normal charges for online searching (actual connect costs plus 15% surcharge for paper, phone, etc.) The average cost to subsidized users was $10.31.

One hundred and fifty-four subsidized searches were performed during the year; evaluation forms were returned for 107 searches (69%). As is true for a number of other studies the independence of responses is not reported, i.e., it is unclear whether each search evaluated represents a unique user, or whether some users had multiple searches and thus returned more than one evaluation. Searches were run on a range of data bases, with heaviest use occurring for Psychological Abstracts Biological Abstracts and ERIC. An average of 1.4 data bases were searched for each request.

The details of the evaluation form are not reported but it appears that a dichotomous choice, satisfied/not-satisfied, was one of the items, since other search parameters are related to this aspect in the report. More than 80% of the searches evaluated were considered to be satisfactory. For 70% of the evaluation forms it was noted that the responding user indicated that she/he would do another computer search in the future.

The authors indicate that the evaluation forms were examined for factors that would account for student satisfaction or dissatisfaction. They conclude that neither the student's status nor purpose in requesting the search were related to student satisfaction. They also indicate no relationship between the cost of the search and expressed satisfaction. It is further suggested that no relationship exists between the percentage of relevant citations retrieved and satisfaction, because some searches in which all citations were relevant did not satisfy the user, while other users were satisfied with less than one relevant citation in ten in their results. Using chi square they indicate a relationship between the number of citations generated and student satisfaction. All students receiving more than 100 offline citations were satisfied. All students receiving at least 40 relevant ciations were satisfied.

The Search Interview

Quantitative studies related to the search interview are far less common than those which consider user evaluations of online searching. Carmon (1975, p.6) reviewed the published literature for reference services and computer-based retrieval services, and concluded that the what of library reference work (and by equivalence data base searching) was somewhat defined, but that the how aspect was largely undefined. Taylor (1968, p.180) called the negotiation aspect of a reference interview a complex act of human communication, with the user trying to describe for another individual not what he knows, but something he does not know. At that time he suggested that quantitative information about the process was non-existent. More recently, Lynch (1977, p. 11) has confirmed that there has been little empirical analysis of the reference interview. Carmon's study (1975) is most related to the current investigation. Other reports, peripherally related, are also discussed in this section.

Taylor's five filter model (1968) of the reference interview was first proposed in the late 60's. He conducted a number of interviews with special librarians and information specialists to determine their methods of question negotiation. He states that this group was selected because they are usually concerned with substantive questions, receive inquiries from motivated and critical users with pre-knowledge of what is acceptable as an answer, and are familiar with negotiation techniques. Based upon information obtained in these interviews he suggests five steps in negotiation, from which the librarian selects data useful in conducting the search. The five steps include:

1. determination of [the] subject; [of the request]
2. objectives and motivation; [of the user]
3. personal characteristics of [the] inquirer;
4. relationship of inquiry description to file organization;
5. anticipated or acceptable answers. (p. 183)

Each step implies that the librarian will assume a question-asking role and that the user's role will focus on information-giving.

Videotape scripts illustrating "poor" techniques of query-negotiation and elements of a "well-negotiated" query are included in Jahoda's instructional module for answering reference questions (1977, Appendix D). The situation simulates reference desk encounters in an academic library. The poorly negotiated information request involves nine questions asked by the librarian. In the well-negotiated request the librarian asks 15 questions.

Somerville (1978) takes a more prescriptive approach to the search interview by including a list of 20 things the search intermediary should do during the interview. She suggests that the searcher:

- Ask questions of the user to ensure your understanding of the subject . . .
- Determine if the user prefers a comprehensive/search or a narrow one . . .
- Make sure that you identify all the restrictions that the user wants placed on the search strategy. (p. 23)

Some correspondence with the five filter model is evident.

Implications and prescriptions for query negotiation have gone beyond the purely mechanistic aspects of the interview. They often include not only what should occur (the transfer of information) but also the manner in which it should occur. For example, Jahoda's list of good and poor ways to negotiate a request suggests that in the good negotiation the librarian will make eye contact with the user, give the user full attention, make the user feel at ease, show empathy for the user, and be aware of non-verbal cues (1977, p. 25). Somerville's list of DO's" notes that the interviewer should make the user comfortable in discussing information needs by utilizing interviewing and counseling techniques (p. 23).

Gothberg (1975, 1976) addressed the non-mechanistic aspect of the reference interview in her study of the effect of immediate and nonimmediate verbal-nonverbal communication behavior by librarians on user satisfaction. Two reference librarians were trained to display immediate or nonimmediate verbal-nonverbal communication during a reference interview in a public library. Immediacy indicates a quality of liking or closeness in an interpersonal relationship. Each librarian was involved in immediate and nonimmediate reference interviews. Reference transactions were recorded and video-taped. Library users (60) were selected on the basis of their availability in the library, need to negotiate the reference question and willingness to answer the questionnaires.

After the interview the investigator approached the user and asked her/him to indicate, by means of a questionnaire, satisfaction with the reference interview, satisfaction with the user's own performance in negotiating the reference question, and satisfaction with the transfer of information. Verbal and non-verbal components of the recorded interviews were coded for immediacy and nonimmediacy.

User satisfaction with the reference interview and their own performance in participating in the reference interview were significant for interviews with immediate communication. However, when the user's satisfaction with the transfer of information was considered there was no significant effect of immediacy in interviews and this satisfaction measure. The author observes that lack of effect in this area may have meant that displays of verbal and nonverbal liking were not sufficient to bring about the trust necessary for users to divulge their true needs and lack of knowledge about the library and its tools.

Gothberg's study is interesting for several reasons. First, it appears that some defined aspects of user satisfaction may relate to the socio-behavioral qualities of the reference interview, while another measure (satisfaction with the transfer of information) may not. Second, the study is of methodological interest because it is one of the few studies isolated which incorporates observations of a real interview situation, quantifies the observations, and assesses the relationship that exists between the quantified observations and a user satisfaction measure. Finally, in its focus on the social-emotional aspects of the interview, it provides a direct contrast to the present study which primarily focuses on a task area of the interview (information-giving and question-asking).

Lynch's examination (1977, 1978) of reference interviews in public libraries is more like the current study of the search interview for it also focused on the mechanistic aspects of the interview. However, user satisfaction was not considered in the study.

Lynch posed eight questions:

1. How often does a reference librarian interview the patron who presents a reference query?
2. Does this frequency vary according to the type of transaction involved?
3. Are interviews more frequent when the librarian is less busy?
4. When an interview does occur, what gross categories or levels of information are sought by the librarian?
5. How often are the questions of the librarian open questions and how often are they closed questions?
6. Does the reference librarian use the secondary questions (probes) used by other interviewers?
7. How does a librarian discover that the query a patron first presented is not the query he/she wants answered?
8. How many primary questions does the librarian ask the patron in an interview? (1977, P. 43)

Reference interviews in four public libraries were recorded. Of the 366 interviews recorded, 309 were ultimately transcribed. Tapes were not transcribed in their entirety. Since question-asking by the librarian was the focus of the study, only questions that the librarian asked were transcribed, along with enough additional material to make them meaningful. Transactions were classified according to their nature, i.e., directional, holdings transactions, substantive or moving. Interviews were considered to occur in those cases in which the librarian asked the patron one or more questions. The investigator developed an 11 category scheme for analyzing holdings' transactions, and a 20 category scheme for substantive transactions. In addition, questions were analyzed on an open/closed scale, and a scale to assess probing questions by the interviewer. Some difficulties in coding were noted. Findings related to questions five, six and eight posed by the investigator are considered here.

Open questions, i.e., those which allow flexibility in user response, were employed infrequently in the interviews. In holdings and substantive interviews they comprise eight precent of the questions asked. Ninety percent of the questions were closed. Two percent were considered to fall into an intermediate category.

The author notes that the sixth question concerning secondary or probing questions by the librarian could not be resolved (1978, p. 32). In coding there was difficulty in establishing whether a true probe occurred (a librarian tries to elicit more fully information about an aspect of the patron's information need), or whether the question by the librarian was actually an attempt to be sure that the patron's original question had been heard correctly.

Primary questions (questions by which the librarian introduces some aspect of the patron's search for information and asks for content new to the interview) were infrequently used in the interviews. Fifty-two percent of the interviews involved only one question. Eighty-nine percent of the interviews had three or fewer questions.

The final study reviewed here (Carmon, 1975) has had a direct influence on some of the aspects of the present investigation. Carmon's study had two goals; first, to collect descriptive and quantitative data concerning the reference process for computer-based literature searching, and second, to develop a model of the user interface for a projected network model of system use.

The study was conducted at two sites, the University of Georgia and UCLA. Each site provided batch-mode searching of a number of data bases. Georgia's system allowed for both current awareness searching (SDI) and also retrospective searching of earlier years on the data bases. UCLA provided current awareness searching only. Evaluative components in the Georgia results do not make a distinction as to whether retrospective or current awareness searches are being considered.

At the University of Georgia interviews were conducted by four reference specialists; these specialists were by job function involved almost exclusively with computerized literature searching. UCLA results originate form searches profiled at a number of institutions in the California system, with the actual computer run being accomplished at UCLA. Searchers in the California system were more likely to have several other job responsibilities in addition to computerized literature searching.

The University of Georgia results cover a period of five months during which 333 users initiated a search request. Approximately two-thirds of the users were graduate students, although some overlap in status categories is noted. At UCLA graduate students comprised almost half the population studied. Twenty-five percent of the interviews at Georgia were recorded, and ultimately 44 of these interviews were transcribed. During the same period UCLA processed almost 150 search requests. Fifty-nine of the UCLA interviews were recorded and transcribed.

User evaluation forms were sent to University of Georgia and UCLA users. Twenty-five percent of the total University of Georgia users returned an evaluation form. For searches with recorded interviews there was a return rate of 77%. At UCLA 52 user evaluation forms were returned.

On a general satisfaction measure (very useful, of some use, of little use, of no use) responding users at both sites indicated that they were satisfied with the results. At Georgia 96% of the searches that were taped were found to be very useful, or of some use; for all the returned evaluation forms 89% of the searches clustered in these categories. At UCLA 91% of the returned evaluation forms indicated that the user found the search to be either of some use or very useful. At both sites the most frequently ranked specific dislike expressed about the service concerned recall, i.e., users ranked no way to judge completeness high on the disliked features list.

Transcribed interviews were coded at each institution. Slightly different, self-developed coding systems were used at each site. Frequencies for events were noted by the occurrence or non-occurence of specific events; these events were not ultimately related to the users expressed satisfaction measures. For both sites it was determined that the events coded were likely to occur throughout the interview. For example, the searcher might focus first on question negotiation, inject some descriptive information about the system, do some strategy development, and return to more question negotiation.

The low frequencies noted for a number of categories in the transcribed interviews were of interest. For example, at Georgia there were no instances in which the searcher asked the user to restate the question, an occurrence which might elicit additional information. Asking the user whether any other concepts should be added to the search occurred in only 10 of the 44 transcribed University of Georgia interviews. The users were asked to give an estimate of the available published literature in nine interviews; relevant author terms were sought from the user in thirteen of the interviews. Language requirements of the user were directly sought in ten interviews; the user was asked to confirm the strategy in only three of the interviews. If, as suggested in previous writings on the interview process, question-asking is an important role for an interviewer, it seemed that an examination of the magnitude of question-asking as related to user information-giving and the ultimate assessment of results would provide an interesting framework for teaching the interview process, and also practical application in actual search interview situations.

III. Methods and Procedures

This section considers methodological and procedural aspects of the study which establish the perspective for reviewing results. Specifically discussed are (a) the research design, (b) limitation of the study to a particular data base, (c) the environmental constraints which constitute the framework from which subjects were selected, (d) data collection instruments employed in the study, (e) data collection procedures, and (f) data treatment methods utilized in generating measures appropriate for describing the results.

Design

The study is characterized as descriptive research focusing on three areas. These include the examination of interrelationships of user assessment areas often included in evaluating schemes applied to operational online search services; the examination of differences in user response styles exhibited by defined subsets of the universe of users who utilize online search services; and the examination of relationships between specified events occurring in the search interview and the user assessment of search results.

Data were collected in a natural setting, that is, collection was incorporated into the normal pattern of procedures employed in responding to a request by a user for an online search. Artificial aspects introduced in this study included asking the user to complete a request form, tape-recording of the search interview, and requesting that certain users participate in an evaluation of search results (questionnaires mailed to users).

Interviews were transcribed, and quantified by means of a particular content analysis scheme (Bales, 1950) applicable to the questions examined in this study.

Online Data Base

This study was limited to interviews and searches conducted for a particular online data base, MEDLINE. Several factors supported this choice. First, the MEDLINE data base has been widely available for searching for a relatively long time (McCarn & Leiter, 1973) in contrast to some of the other online data bases available through the commercial vendors. Because of this, it offered the opportunity for investigation in an atmosphere uncomplicated by variables associated with the use of more recently available data bases (e.g., few users available for study because they are as yet unaware of a new service; searchers interviewing for, and searching on a data base which is unfamiliar to them). Second, because of its broad subject content MEDLINE is often applicable to requests in a number of interdisciplinary areas (medicine, biology, psychology, and some social sciences). Interviews for MEDLINE may mirror to some extent the interview employed in searching other data bases.

Third, certain aspects of question negotiation for MEDLINE searching (thesaurus use, strategy development, establishment of preferences for time coverage and citation format, etc.) were likely to provide a common denominator in interviews at all of the search sites; furthermore, these aspects are comparable to those utilized in question negotiation for several other data bases. Finally, an earlier experimental study of relevance judgments focused on assessments made by groups of individuals in the biomedical field (Rees & Schultz, 1968) and established, within this context, a basis for believing that depth of research knowledgeability may be associated with more stringent relevance assessments. An examination of this aspect in an operational situation was of interest in the current study.

Thus, the limitation to an examination of MEDLINE searches is not to be viewed as a specific attempt to evaluate the MEDLINE system, but rather MEDLINE was chosen for study because it is a data base which has many features common to other data bases, and MEDLINE users were likely to be similar to users examined in an earlier study.

Sampling Plan

Subjects were in-person users and searchers at three institutions providing MEDLINE searches. The following sections detail characteristics of sites, searchers, and users which define the sampling framework.

Sites. Collection sites included the Vera Shiffman Medical Library at Wayne State University in Detroit; the Medical Center Library at the University of Michigan in Ann Arbor; and the Houston Academy of Medicine-Texas Medical Center Library. The first two libraries were included as participants because they had expressed interest in cooperating in the study, and their locale was convenient for regular visits by the investigator to pick up tapes and other materials. The third library was included when data collection appeared to be progressing at a rate slower than anticipated. Materials from Houston were mailed to the investigator. All institutions, searchers and users were promised anonymity in the reporting of final results.

The Shiffman Medical Library serves both the immediate population of faculty, students and researchers associated with Wayne State University, and a more diverse group of urban and hospital-related users with information needs related to the health care field. Shiffman is also the Regional Resource Library for Kentucky, Ohio and Michigan (KOM).

The Medical Center Library at the University of Michigan is located in the hospital complex of the University Medical Center.

The Library serves the information needs of researchers, faculty, students, and clinicians affiliated with the Medical School and other health profession schools of the University. It also serves as a resource for individuals associated with a number of prominent research institutes at the University.

The Houston Academy of Medicine-Texas Medical Center Library is a consortium library governed by representatives from major participating institutions -- the Houston Academy of Medicine, Baylor College of Medicine, Texas Women's University, the University of Texas at Houston, and Texas Medical Center, Inc. In addition, it is the Regional Resource Library for the South Central Regional Medical Library Program.

All of the libraries charge for MEDLINE searches, with basic fees ranging from $5 to $15 or more, depending on the number of backfiles searched and the format of the printouts (e.g., inclusion of abstracts). One library prefers to have the user present for the interview and terminal search, another library intermixes these procedures (some interviews are combined with the search, some searches are run after the user leaves), and the third library, in most cases, separates the two procedures (the search is run sometime after the interview with the user). However, some terminal searches with the user present were noted as occurring in all of the libraries for this particular study.

Searchers. Eleven searchers participated in the study. All searchers can be considered to be relatively familiar with MEDLINE since they had conducted 125 or more searches in the 12-month period prior to the study. One searcher had at least a year's experience with MEDLINE; all others had two or more years of experience. Undergraduate majors of searchers were predominantly in areas other than the physical or life sciences. Two searchers had undergraduate degrees in the sciences.

Searchers indicated that training for MEDLINE searching had been accomplished in a variety of ways. Four searchers had attended training sessions of varying lengths at the National Library of Medicine; the others were trained for MEDLINE searching through several mechanisms (singly or in combination). These included NLM sessions scheduled in their vicinity, use of the MEDLEARN training sequence, training by a more experienced colleague, and self-teaching.

Users. Several limiting factors, either external to the study or design imposed, determined the sample framework of the potential pool of users investigated in this study.

The pool from which users could be drawn included all unique, in-person users requesting a MEDLINE search at one of the search sites during the data collection period. Thus excluded from the outset were search requests received by mail or phone, search requests from a user who had already participated in an earlier session, and search requests from users who asked for another data base even though MEDLINE may have been subsequently used to satisfy the information need (all of the sites provided a range of search services on data bases other than MEDLINE).

A second limiting factor on the potential pool of users was self-selection of individuals for participation in the study. In recognition of concerns for user and searcher privacy in regard to tape-recording of interviews, and in compliance with guidelines from several committees charged with monitoring all studies which involve human experimentation, a consent form for participants was developed (Appendix A). Searchers were asked to note the reason for not recording all searches they conducted during the data collection period. Known decisions not to participate (either searcher or user) occurred in a ratio of less than 1:2 when considered against decisions to participate. The ratio may be higher since less than complete information was available from one data site. In several instances searchers noted that the user did not wish to participate because the user was "in a hurry" and did not wish to take the time to complete the request form. In almost half the non-participating occurrences the decision not to record was searcher initiated and appeared, in a number of instances, to be related to lack of immediate access to the tape recorder or general lack of time to complete all the procedures. Most instances of lack of time for procedures occurred at one site which had a relatively high volume of searches during the data collection period.

A third, design-imposed factor was the ultimate selection of participants to receive evaluation forms. Subjects receiving evaluation forms were selected from the total group of recorded participants according to design criteria, i.e., subjects selected had indicated that their primary status was either faculty (instructor, assistant professor, associate professor, professor) or student (undergraduate, professional school student, or graduate student). Furthermore, subjects indicating either of the previous status conditions had to have also indicated that the search was primarily for their own use. This was to preclude a situation in which the ultimate evaluator of the search was someone other than the person taking part in the interview.

Faculty subjects selected to receive evaluation forms were further restricted in that their stated purpose in obtaining the search had to be research-related (grant project, preparing an article or review for publication, on-going research leading to publication). This restriction was imposed to parallel the conditions noted by Rees and Schultz

(1967) in their experimental study of relevance in which it was noted that knowledgeability and involvement in research were related to more stringent relevance assessments. In other words, it was not desired that faculty searches of a more sporadic or superficial nature (e.g., search for immediate clinical application, background material for a speech, material for class instruction) be included in the evaluation.

A final situational restriction was imposed after-the-fact. In a few cases two users participated in the interview. Since the search interview is considered to be primarily a dyadic communication process, these interviews would be misleading if analyzed for information-giving behavior by the user. Furthermore, they would present problems in determining the individual most appropriate for evaluating the searches. Searches of this type were excluded.

Data Collection Instruments

Data collecting instruments for this study included a user request form for collecting background information concerning the user, tape recordings of interviews, user evaluation forms sent to participants meeting study criteria, and coded transcripts of selected interviews.

User Request Form. The user request form (Appendix B) was developed for two purposes. First, to obtain primary information about each user which allowed a decision to include or exclude a particular user from the evaluation process. Questions 7, 10, and 11 (ultimate user of the search, status of requester, and purpose of the search) fulfill this function. Second, for comparative purposes it was also desirable to more fully describe background characteristics of the user that might clarify the framework from within which the user made an evaluative decision. Questions 6, 12, 13 and 15 (number of previous online interviews, degrees completed, familiarity with and sustaining interest in the search topic, and ability to list relevant citations) were of this type. Question 14 (narrative search statement) was included for comparison with an earlier study (Carmon, 1975) in which it was noted that users supply fairly brief statements concerning their search topic. Other questions on the form were designed to accommodate information usually included by one or all of the search sites in their already existing forms.

Tape Recordings. Tape recordings were utilized to capture all verbal interaction between searchers and users. They were turned on as soon as the consent form was signed and turned off at the completion of the interview.

User Evaluation Form. A standardized, generally accepted instrument for eliciting user evaluations of searches does not exist.

The form used in this study is self-developed, but it reflects closely areas considered in other evaluation schemes (Daniels, 1978; Hitchingham, Note 1).

Specifically, the evaluation form (Appendix C) was designed to elicit information from the user in five areas: (a) the relevance score for the search, where relevance is determined as the number of relevant citations indicated by the user when she/he examines the search results, divided by the number of citations retrieved; (b) the user's satisfaction with the proportion of relevant citations retrieved; (c) the user's concern for recall in considering the search results; (d) the user's assessment of the value of the search in meeting the need prompting the search request; and (e) the user's perception of the searcher in the interview process.

In cases where the user estimated relevance, the upper level score of the estimate was used (e.g., if 51% to 60% was marked, 60% is the value). Concern for recall and user perception of the searcher in the interview are summative measures.

The concern for recall score is the average (mean) score obtained by summing user responses to the four statements relating to recall on the user evaluation form, and dividing this sum by the number of statements responded to (three or more). The statements assess whether the user (a) believes that the search retrieved most of the relevant citations in the data base, (b) is concerned because there is no way to judge completeness, (c) is concerned because the results omitted relevant citations known to the user prior to the search, and (d) considers that fewer relevant citations than expected were retrieved. Scale values noted by the user for the second, third, and fourth statements were subtracted from eleven to conform to the order of the first statement. Lower scores indicate more concern for recall, higher scores indicate less concern for recall.

The user perception of the searcher in the interview is the average (mean) score obtained by summing user responses to five statements concerning the searcher on the user evaluation form, and dividing the sum by the number of statements responded to (four or more). The statements assess, on a disagree (one) to agree (ten) scale, whether the user (a) believes that the searcher was knowledgeable concerning the use of the data base for the question, (b) feels that the searcher understood the request, (c) believes that the searcher understood the user's purpose in initiating the request, (d) believes that the searcher was thorough in exploring all aspects of the search question, and (e) believes that the searcher suggested terms appropriate to the subject of the request. Lower scores indicate lesser agreement to the statements concerning the interview.

The instrument can be considered to be valid insofar as it reflects common elements noted in other evaluations forms. Segments of the form which employ summative measures (concern for recall, perception of the interview) appear to be sufficiently reliable when item consistency is considered. Reliability coefficients (alpha) of .76 for the concern for recall, and .90 for the perception of the interview were established (Nunnaly, 1967, p. 196).

The major departure from most of the other evaluation forms reviewed was the use of a 10-point rating scheme for most of the questions instead of the three or four-point scale more commonly used. Application of expanded scales is consistent with general scaling theory (Nunnaly, 1967, p. 521).

Coded Transcripts. Transcripts of selected interviews were unitized by the investigator and coded by trained judges. The categories applied to the transcripts were those defined by Bales in Interaction Process Analysis (1950). The scheme is outlined in Figure 1.

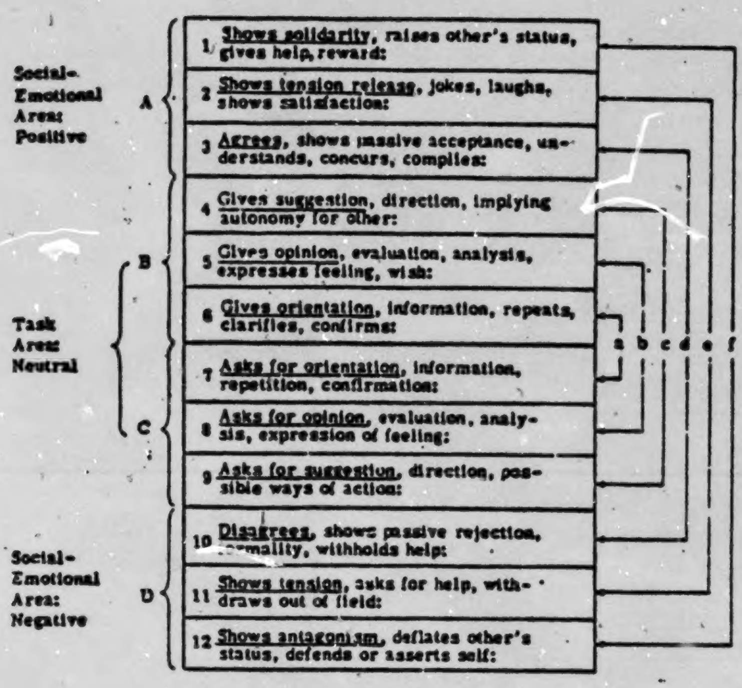
Interaction Process Analysis (IPA) is "an observational method for the study of social and emotional behavior of individuals in small groups"(Bales, 1968, p. 465). The method centers on the coding of interaction units (verbal and non-verbal) occurring when small groups are involved in a problem-solving task.

An interaction unit is "the smallest discriminable segment of verbal or non-verbal behavior to which the observer, using the present set of categories after appropriate training, can assign a classification" (Bales, 1950, p. 37). In this study the interaction unit is defined as the smallest discriminable segment of verbal behavior.

Units are coded according to the 12 category system. The first three categories are considered to represent positive, social emotional areas; the fourth through ninth categories represent neutral, task areas; and the tenth through twelfth categories are designated as negative, social-emotional areas.

The system is general in nature and applicable to small-group interaction. Although the system covers the gamut of interactive events that could occur in an interview, albeit in a very broad sense, the primary categories of interest in this study are those noted as Task Area: Neutral in Figure 1, the information-giving and question-asking categories.

In a transcribed interview user information-giving activity is the sum of user units coded in categories four, five, and six (gives suggestion, gives opinion, gives orientation). Searcher question-asking activity is the sum of searcher units, from a transcribed interview, which are coded in categories seven, eight, and nine (asks for orientation, asks for opinion, asks for suggestions).


Social-Emotional Area: Positive
A
B
Task Area: Neutral
C
Social-Emotional Area: Negative
D
a b c d e f

1 Shows solidarity, raises other's status, gives help, reward:

2 Shows tension release, jokes, laughs, shows satisfaction:

3 Agrees, shows passive acceptance, understands, concurs, complies:

4 Gives suggestion, direction, implying autonomy for other:

5 Gives opinion, evaluation, analysis, expresses feeling, wish:

6 Gives orientation, information, repeats, clarifies, confirms:

7 Asks for orientation, information, repetition, confirmation:

8 Asks for opinion, evaluation, analysis, expression of feeling:

9 Asks for suggestion, direction, possible ways of action:

10 Disagrees, shows passive rejection, formality, withholds help:

11 Shows tension, asks for help, withdraws out of field:

12 Shows antagonism, deflates other's status, defends or asserts self:

KEY:

a Problems of Communication
b Problems of Evaluation
c Problems of Control
d Problems of Decision
e Problems of Tension Reduction
f Problems of Reintegration

A Positive Reactions
B Attempted Answers
C Questions
D Negative Reactions

Figure 1. Categories in the Interaction Process Analysis Scheme.

Note: Reprinted from Interaction Process Analysis by R. F. Bales, 1950, p. 9, by permission of the University of Chicago Press, Copyright 1949 by the University of Chicago.

IPA appeared appropriate for this investigation because (a) it is an established method of content analysis (Borgatta and Crowther, 1965; Heyns & Lippit, 1954; Holsti, 1969) (b) definitions for scoring the categories were available (Bales, 1950, pp. 177-195); (c) categories four through nine were directly related to areas being investigated; (d) "norms" in the sense of profiles for other small-group interactions existed (Bales & Hare, 1965); and (e) the system is applicable for future comparative studies of search interviews.

Data Collection

In this section preliminary procedures, data collection period, site procedures, collection of evaluation forms, selection of tapes for transcription, and coding of transcribed tapes are described.

Preliminary Procedures. Prior to initiation of data collection a meeting was held to familiarize local participating searchers with the procedures and tape-recording process. Out-of-state searchers were mailed the procedures (Appendix D), forms, and blank tapes. At this same meeting comments were received concerning the user request form and the user evaluation form; some revisions were made based upon the input received. Searchers were aware that a study of the tape-recorded interviews constituted a part of the research plan, but they were not aware of the specific methodology employed (i.e., the focus on information-giving and question-asking).

Data Collection Period. Site data was collected over a period of 4-1/2 months, from mid-June through the end of October 1978. Collection periods for each site were not concurrent, since final approval of the participant consent form occurred within different time frames. Two sites collected data for approximately 3-1/2 months, one site collected data for a lesser period.

User evaluation forms were accepted through mid-December 1978. The sequence of initial contact by mail, with two mail follow-ups and a telephoned reminder, required almost two months for completion after the end of site data collection.

Site Data Collection. Users requesting a MEDLINE search at the sites were given a consent form which briefly outlined the project plan. If the user agreed to participate and signed the form tape-recording began. Interviews were then conducted in the usual manner.

Several searchers commented that they were quite aware of the recording process. However, since there was no preknowledge of what was being examined in the interviews, this does not appear to be an influencing factor in the ultimate consideration of the interviews. When the tapes were later reviewed by the investigator, no artificiality in interviewing procedures was readily apparent.

At the end of the interview the tape recorder was turned off. If online citations were given to the user the number of citations was noted. If the total number of offline citations was known at the time, the number was also noted. In most cases the total number of citations was not immediately known, since backfiles were also searched. For these cases citation counts were recorded when the results were received.

The investigator periodically visited the two local sites to pick up tapes, consent forms, and request forms. The non-local site mailed materials on a weekly basis.

During the data collection period a total of 132 interviews were recorded at the three sites. Four of these interviews did not result in a MEDLINE search because after discussion with the searcher the user was directed to an alternative information source; two interviews were recorded but the request forms were not available.

Of the remaining 126 unique users, 65 were determined to be inappropriate to this study. Almost one-third of this number was excluded because the search results were not intended for the person present at the interview; thirty-five users were excluded because they indicated a status outside the scope of this study (e.g., attorneys, librarians, intern-residents). Four of the remaining users were excluded even though they were faculty because they indicated a non-research purpose in requesting the search; five were omitted because more than one person was interacting on the tape. The residual pool of appropriate users included 23 faculty members and 38 students (N = 61).

Although the recording procedures led to recording of more interviews than would ultimately be utilized, this seemed appropriate. Interaction which would eliminate users at the outset (e.g., determination of status, purpose, etc.) is considered in Taylor's five-filter model (1968) to be an integral part of the interview process.

User Evaluations. All appropriate users were sent a cover letter (Appendix E) explaining the project and an evaluation form. A return envelope was enclosed. If the evaluation form was not returned after 2-1/2 weeks, a hand-written note was sent to the user prompting her/him to return the form. If, after another two weeks, the form was still missing a third note was sent. As a final step an attempt was made to contact the user by phone and urge the return of the form.

A total of 55 forms were returned (90%); however, one of the forms was returned several months after the cut-off date and was not included in the analysis. For the 54 included searches 22 were from faculty (96% return) and 32 were student searches (84% return).

Transcription of Tapes. Although it was originally intended that interviews for transcription would be randomly selected from the pool of cases with completed evaluation forms and good recordings, this did not prove feasible. Preliminary work with the tapes indicated that they would be more difficult to transcribe than had been anticipated from an earlier report of similar transcription (Carmon, 1975). A more recent study involving taping of interviews in public libraries (Lynch, 1978) suggests that transcription is a lengthy, complex process. In the present study some familiarity with MEDLINE procedures and medical terminology proved to be necessary for an accurate transcription of the tapes. For this reason the investigator transcribed all tapes rather than delegating them to clerical personnel as originally intended.

Even though the investigator was more familiar with the context and terminology involved, transcription proved to be a time-consuming task, involving in several instances up to 20 hours per interview. This was particularly true for those interviews which involved use of the terminal while the user was present. Dialogue had to be detected against the background noise of the operating terminal. Knowledge of the length of time likely to be involved in transcribing each interview indicated that a delay for random selection subsequent to the return of all evaluation forms was not appropriate within the time-frame of this study.

Interviews were selected for transcription by two criteria: first, to provide diversity in the searchers and user types included, and second, to allow transcription to occur along an extended period. Interviews thus selected were those associated with the first faculty and first student evaluation forms returned for each searcher. This selection does not necessarily reflect the initial interviews recorded by searchers. Some searchers had recorded several inappropriate users before encountering a request which met study specifications. Other searchers recorded an appropriate user early in the process, and were involved with an alternate appropriate user much later on. Two searchers had only student interviews available for transcription. A total of 18 interviews were transcribed (10 students, 8 faculty). One searcher had conducted relatively few interviews, none were appropriate for transcription.

The interviews were divided into interaction units according to Bale's methodology (1950, p. 37) at the time of transcription.

Coding of Interviews. Transcribed interviews were coded by two experienced searchers at the Computer Search Service, MIT Libraries. Coders were familiar with MEDLINE and other online data bases, and as full-time searchers were involved with interviewing on a daily basis. It was believed that this familiarity with the interview process would provide a good background for interpreting the unitized transcripts in accordance with the scoring system.

The coders were first sent copies of the definitions of categories (Bales, 1950, pp. 177-195) and copies of several transcripts obtained from the University of Georgia study (Carmon, 1975), so that they might become familiar with the coding scheme. Once the coders had some practice at applying the scheme to interviews an extended training session was scheduled.

At the end of the training session coders were able to demonstrate 82% agreement when scoring three pages containing a total of 96 units (categories 4, 5, and 6 were collapesed, as were categories 7, 8, and 9 since data manipulation would involve total information-giving by the user).

Transcribed interviews for this study were divided between the coders so that each received an approximtely equal number of pages to score. A duplicate copy of one search was sent to each coder so that a reliability coefficient for coding material involved in this study could be established. Three pages were selected at random from the search for the determination. A more stringent method (Cohen, 1960) than simple percent agreement was applied (see Appendix F). A coefficient of .77 was obtained; this appeared adequate for the applications involved in this study which focused on the ranking of results.

User information-giving and searcher question-asking scores for the analyzed searches were obtained by counting the frequency of events coded by category for the searcher and the user. Topics of the 18 transcribed and coded searches are listed in Appendix G. Appendix H includes an example of a coded interview.

Data Treatment

Distribution free tests were used in analyzing the primary questions of this study. Specifically, interrelationships of assessment parameters (value with relevance, satisfaction with the proportion relevant, concern for recall, and user perception of the searcher in the interview) were examined by means of the Kendall rank correlation coefficient. The relationship of user information-giving to relevance and concern for recall scores was also examined in this manner. The examination of searcher question-asking as a contributory factor in user assessments of relevance and concern for recall was examined by means of the Kendall partial rank correlation coefficient.

User differences (faculty, students) were examined by means of the Mann-Whitney U test. Earlier discussion suggested more stringent assessments by faculty in the areas of relevance, concern for recall and perception of the searcher in the interview, and greater information-giving activity by faculty in the search interview; a directional (one-tailed) test was used for these considerations. A non-directional test (two-tailed) was employed to examine differences in satisfaction with the proportion of relevant citations and value.

Version 7 of the SPSS program was utilized for computations. (Nie & Hull, 1977).

IV. Results

Data resulting from this study are considered and analyzed in four sections. The first section describes the data sources. The second section reports overall characteristics of recorded users. The third section addresses the primary questions upon which the study focused, and interprets the results in light of other findings related to the scope of this study. The last section examines the interview profiles developed in the coding process, and discusses the profiles in relation to this study and other "norms" reported for interaction profiling.

Data Sources

General variables considered are available from 126 user request forms, and the 54 evaluation forms from the 61 users selected as appropriate for follow-up evaluations. Respondents included 22 faculty users and 32 student users. One student did not indicate the number of relevant citations retrieved,thus the number of cases for any testing which includes relevance is 53.

Interviews from 18 of the 54 responding users were transcribed and analyzed. This subset includes ten student users interacting with ten different searchers, and eight faculty users interacting with eight different searchers. Factors relating to the analyzed interviews were compared with non-analyzed cases by means of the Mann-Whitney U test. Analysis ($P = .05$) indicated that there was no difference in the distribution patterns for the two groups on the number of relevant citations retrieved, the relevance scores, the satisfaction with the proportion of relevant citations, the concern for recall scores, value scores, and the perception of the searcher in the interview. Thus it appears possible to assume that the searches selected for coding were valid representations of the entire sample studied, insofar as these measures are concerned.

Overall Characteristics

Although online services have been available for several years, more than half (54%) of the recorded users were novice participants in an online interview situation. Slightly more than a third (34%) of the users indicated participation in one to five such interviews; 10.3% of the users indicated they had been involved in six or more interviews (two users did not respond to this question). The percentage of novice users in this study is greater than that noted by Tagliacozzo (47%) in the 1973 study of early MEDLINE users (1975, p. 295), but less than the 70% new users examined by Carmon (1975) in the Georgia/UCLA study.

As noted previously a significant number of interviews (17.5%) occurred in situations where the results of the search were not primarily intended for the individual interacting with the searcher.

This aspect (second-hand information at the interview) has not been directly examined in this study, but if the proportion of such interactions are similarly high in other online operations, an examination of user satisfaction in relation to this factor might be of interest.

Since it was believed that some users might consider more than one status classification on the request form to be representative of their situation, users were asked to designate the one status relating most to their request for a search. Approximately one third (34.9%) of the users designated student status (undergraduate, graduate, medical, dental or law). Twenty-eight users (22.2%) indicated faculty ranks (instructor, assistant professor, associate professor, professor); almost half of this group were assistant professors. Other medical personnel (residents, physicians, nurses, dieticians, therapists, pharmacists) accounted for 18.3% of the MEDLINE users. Three individuals indicated post doctoral appointments and ten users designated their status as academic researcher. This latter category however, was often used by individuals who had also indicated that the search was not primarily for their own use, suggesting that this category is likely to be chosen by individuals in supporting functions (e.g., lab assistants). Thirteen users were less directly involved with the medical field (attorneys or legal assistants, librarians, professional researcher, administrator, biomedical engineer). In five cases the status of the user was unclear. All of these cases involved searches for a different end user than the person being interviewed.

Whatever their status, many of the MEDLINE users indicated that their primary reason for requesting a search was research. Almost 64% of the users indicated a grant project, preparation of an article for publication, or ongoing research leading to publication as the purpose of their MEDLINE request. If dissertations are included 72% of the users can be considered to have indicated research intentions. Other academic related reasons included material for instruction, masters thesis, term papers or class project (14.2%). Immediate clinical applications, where this is defined as use for the diagnosis or treatment of a patient or client, accounted for only 2.4% of the searches. The remaining purposes included seminar or talks, legal preparation, curiosity, and compilation of a bibliography with no immediate application.

In an earlier study of MEDLINE users from several institutions (Taglacozzo, 1975, p. 296) 38% of the users clustered in what was designated as a lower degree category (H.S., B.A., B.S., M.A., M.S.). In this study 54% of the users indicated these degree levels. Twenty-three percent of the users are included in a second tier of educational attainment (medical, dental or law degree, two Master's degrees or M.D. plus Masters). Eighteen percent of the users had Ph.D. degrees.

For 5% of the cases degree status was either unclear or missing.

Although many of the MEDLINE users (68%) indicated that they were already aware of some recent publications related to the topic of their search request, only 20% of the users listed one or more specific citations on that part of the request form which asked for bibliographic information concerning known relevant citations. Fifty-three percent of the users said that they were regular readers of journals which had articles related to the topic of their request. A relatively small number of users (13.5%) had already published an article on the topic of their search request. Most of the users (87%) indicated a continuing interest in the area of investigation.

In written descriptions of their search needs (the narrative statement portion of the search request form) many MEDLINE users were not particularly communicative. For the 120 users completing this section, statements ranged from 2 to 76 words ($\bar{X} = 22.3$, SD = 16.3). Carmon (1975, p. 33) noted an average of 28.6 words in request statements. When commonly occurring verbs, prepositions, conjunctions, etc. were eliminated in that study, an average of 7.2 information conveying words remained. Whatever the reasons, it seems that users are unlikely to reveal all of the parameters and specifications of their search request by means of the written search statement. If this is true, clarification of the request in the search interview would appear to be of considerable importance.

Search Results

Earlier these questions were posed:

1. How do specific user evaluations in the areas of relevance, satisfaction with the proportion of relevant citations retrieved, concern for recall, and perception of the searcher in the interview related to the user's indication of the value of the search results in meeting the need prompting the search request?
2. Do students and faculty MEDLINE users differ in evaluative designations, and in information-giving during the interview?
3. What are the interrelationships between user information-giving, searcher question-asking and two traditional foci of assessment -- relevance and recall?

This section considers the results associated with these questions.

Interrelationships of User Assessments

When the evaluation scores are ranked and compared to the ranked value scores by means of the Kendall rank correlation coefficient, significant correlations are noted in each instance. Table 1 indicates, in decreasing order of relationship, the tau values observed for the cases.

TABLE 1

Correlations of Evaluative Measures
With User Designations of Value

Measure	Kendall's tau
Satisfaction with the proportion of relevant citations retrieved	.66*
Concern for recall	.60*
Relevance	.47*
Search interview (User perception of Searcher)	.42*

* p = .001

Like users in most other reports of online evaluations, MEDLINE users, when considering the value of the search results in meeting the need prompting their request, indicated generally favorable responses concerning value. This was true even though users were provided a ten-point scale to indicate value rather than the more customary three- or four-point scale. No users marked the lowest designation (no value) and almost 60% of the users indicated a value of eight or higher. Figure 2 outlines the distribution of responses on the value scale. Application of an extended scale does, however, appear to have some merit in increasing the variation of response patterns, which is particularly useful if comparisons of magnitude, as in this study, are desired.

The favorable response pattern on value is coupled with a fairly low pattern of relevance scores. The mean relevance score for all searches is 45%. Thirty-one of the searches had relevance

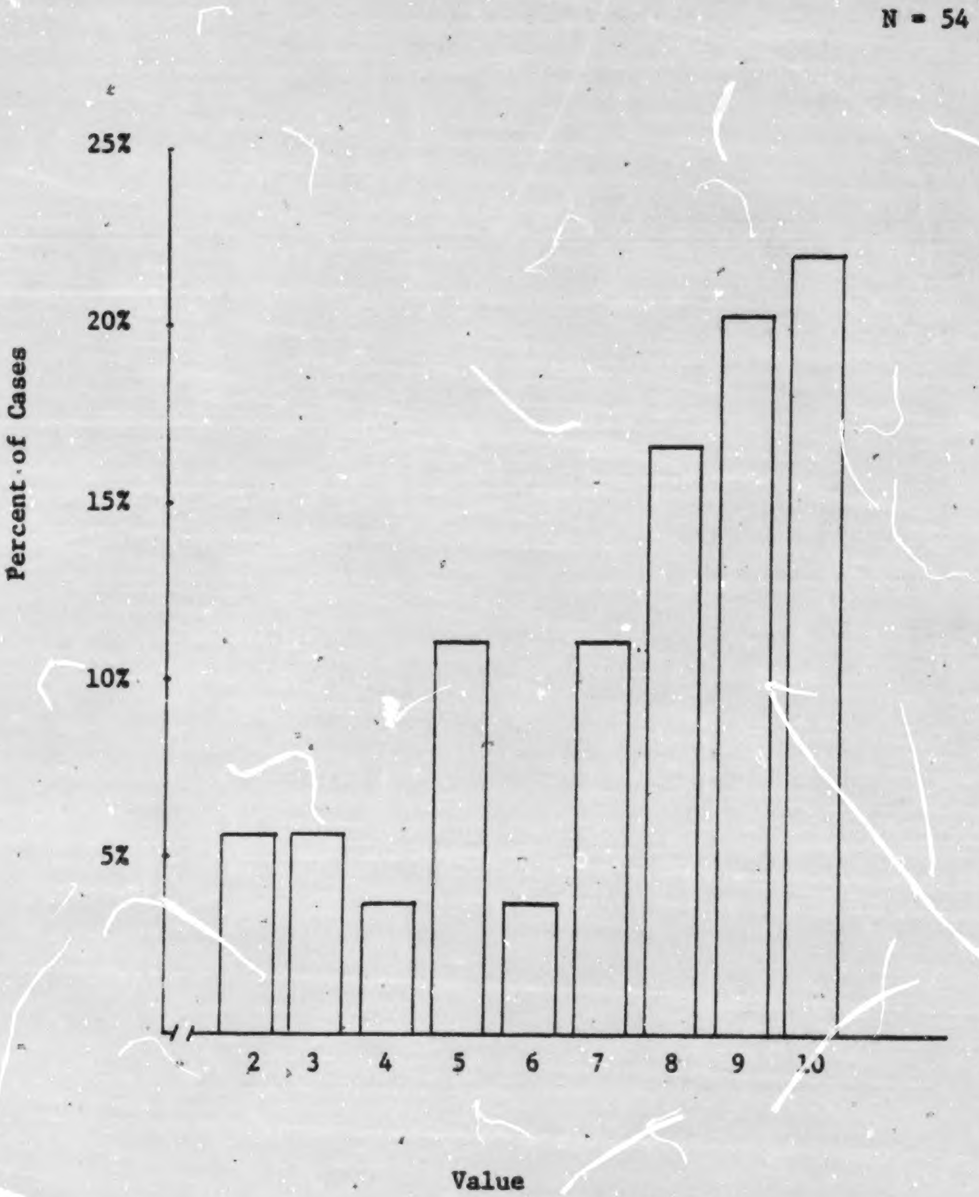


Figure 2. Value of MEDLINE Search Results in Meeting the Need Prompting the Search Request.

scores of 50% or less (i.e., in these searches half or fewer of the citations sent to the user were considered to be relevant to the search question). Figure 3 outlines the distribution of relevance scores. The distribution tends to be U-shaped, a pattern which was noted by Rees and Schultz (1967, Vol. 1, p. 118) when they pooled relevance ratings for all documents in their experimental study of relevance.

As noted previously, relevance and value scores show only a moderate relationship to each other ($\tau = .47$). Tagliacozzo (1977) has suggested that dissonance in user responses may occur; in that study dissonance (contrary responses) were noted for helpfulness and usefulness. In the present study dissonant responses appear to be concentrated in cases with lower relevance scores. For example, all cases of medium (5 to 7) to high (8 to 10) value designations are associated with medium to high relevance scores, but lower relevance scores (0-33%) are associated with low to medium, and high value designations. Relevance scores are simply ratios determined by dividing the number of relevant citations by the total number of citations retrieved, and do not reflect the actual number of "good" citations the user gets from the search. For example, a user might find four relevant citations out of a total of five retrieved; the relevance score is 80%. Another user might find 40 relevant citations in a search which results in a total of 80 citations; the relevance score in this case is only 50% The second user, however, has more relevant citations. It has been suggested that searches that produce more relevant references are associated with helpfulness (Tagliacozzo, 1977, p. 246) or satisfaction (Kobelski & Trumbore, 1978, p. 16). In the first case searches with six or more useful references were contrasted with those achieving five or fewer on a dichotomized helpfulness scale. In the second case the authors noted that all users receiving 40 or more relevant citations were satisfied. They also indicate that all users receiving total outputs of 100 or more citations were satisfied.

In the present study 68% of the users received 40 or fewer relevant citations. Figure 4 indicates the distribution of relevant citations. By sheer volume of relevant citations, the results in the current study appear to be "better" than those indicated in the public library experiment (Mick, 1977). In that study 60% of the second year users received 0 to 15 relevant citations, in the third year 75% of the users received 0 to 15 relevant citations. Yet it can be recalled that 68% to 76% of those users indicated that their search was of major or considerable value. This suggests that the actual number of relevant citations received by the user may not be of paramount importance in value estimations. In the current study of MEDLINE users, when the number of relevant citations received are ranked and compared to ranked user value designations a moderate relationship ($\tau = .53$) similar to that noted for relevance is observed. A negligible relationship ($\tau = .16$)

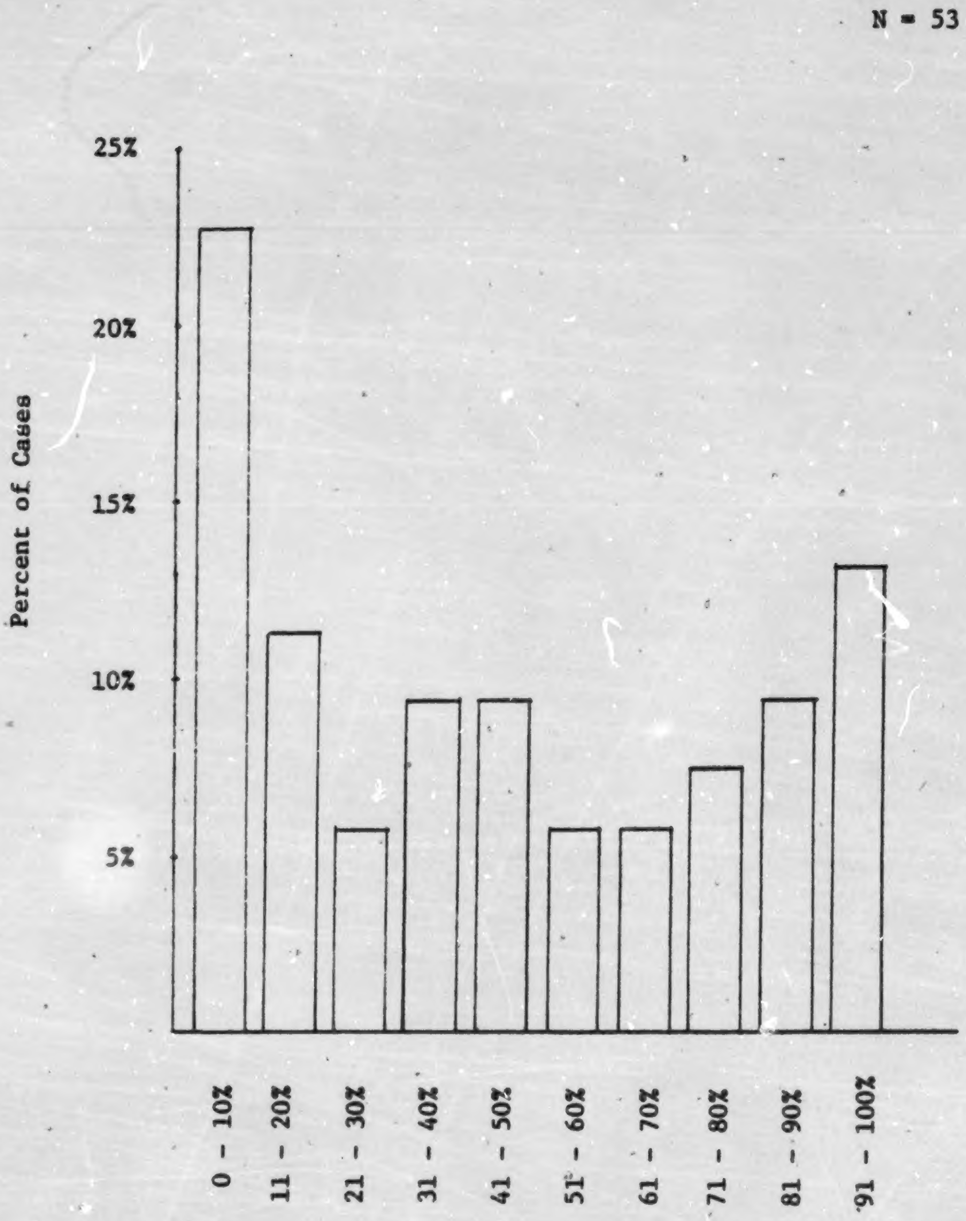


Figure 3. Distribution of Relevance Scores for MEDLINE Searches.

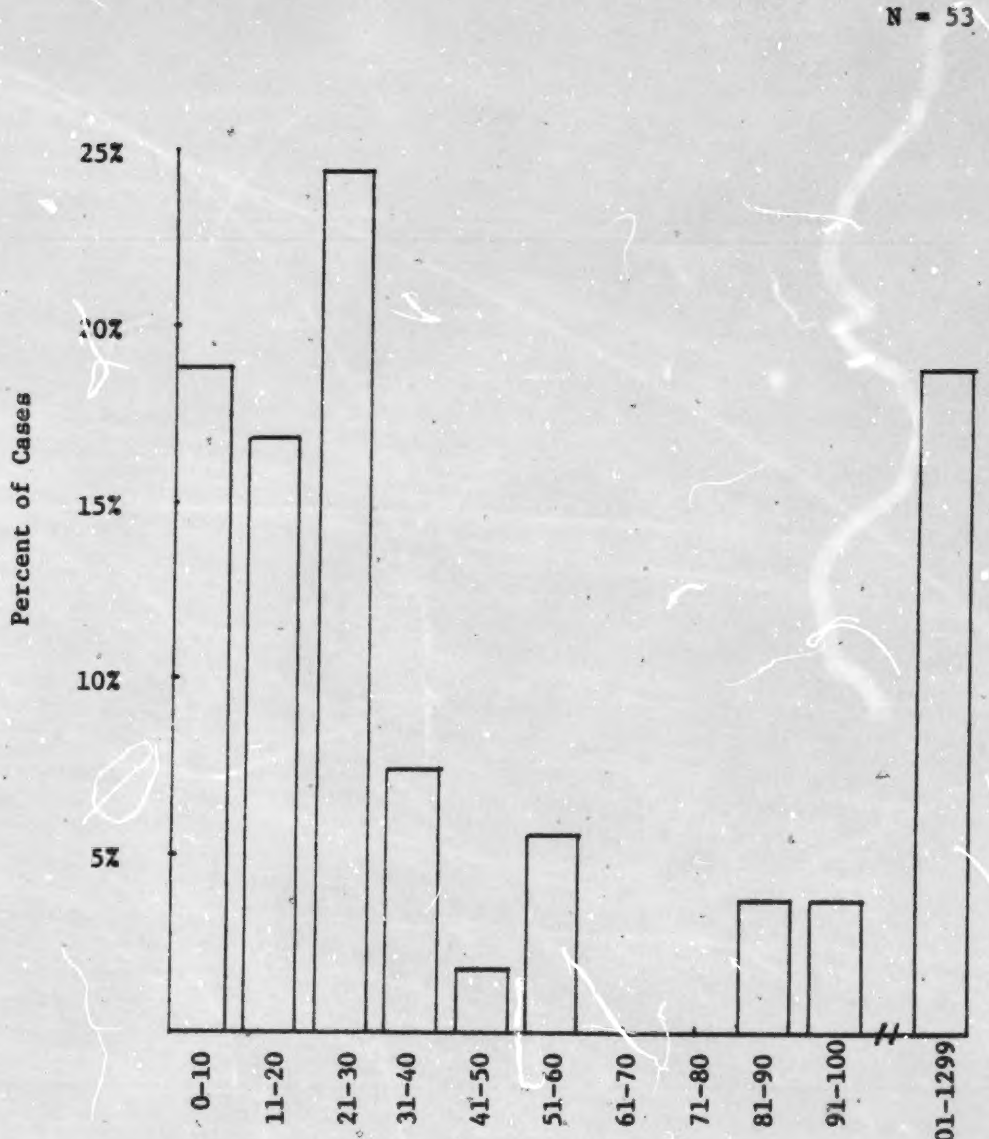


Figure 4. Distribution of Relevant Citations in the MEDLINE Searches.

appears to exist between the total number of citations retrieved and value.

The obtained search interview score (the mean score computed for five questions to the user about the searcher's role in the interview) showed the least relationship to the value score ($\tau = .42$). User impressions of the searcher in the interview situation were generally favorable. Scores of 7 or higher were observed for 77% of the searches. Figure 5 indicates search interview scores for the MEDLINE searchers. Findings reported here do not appear to be inconsistent with those reported earlier by Gothberg (1975) concerning user perceptions about two types of interview situations and user satisfaction with the transfer of information. Users exposed to immediate interviews (a sense of "liking" is conveyed) were more satisfied with the interview itself and better satisfied with their own performance in the interview than were users exposed to non-immediate interviews. However, the two user groups exhibited no difference in their expressed satisfaction with the transfer of information. Although the conditions are different, one might anticipate from the earlier study that the user perception of the interview (a social interaction) will exhibit a somewhat lesser relationship with a judgment of value than other areas more closely aligned with the product (the search results) being evaluated.

Most closely associated with the value designations indicated by the user were the scores on satisfaction with the proportion of relevant citations retrieved ($\tau = .66$) and the mean score, concern for recall ($\tau = .60$). Some of the dissonance noted at the bottom third of the relevance scale (high, medium and low value scores) may be related to a perception by some users that lower relevance means that they have missed some items. Eight of the nine users indicating a satisfaction with the proportion relevant score of three or less had relevance scores lower than 30%. Five of these eight had concern for recall scores less than three.

In the Georgia/UCLA study (Carmon, 1975), users ranked the lack of ability to judge completeness as the most disliked feature of the search services. In the current study the concern for recall score exhibits a considerable relationship with the value score. Repetition in this study of a situation in which users appear to attach some importance to the recall aspect of a search is particularly interesting in light of Cooper's suggestion (1973) that recall may be an inappropriate measure for consideration in retrieval evaluations, i.e., he does not consider it important to the user because, in a general sense, the user cannot "know" that he has missed citations in the results. Figure 6 indicates concern for recall scores. Figure 7 indicates the satisfaction with the proportion of relevant citations.

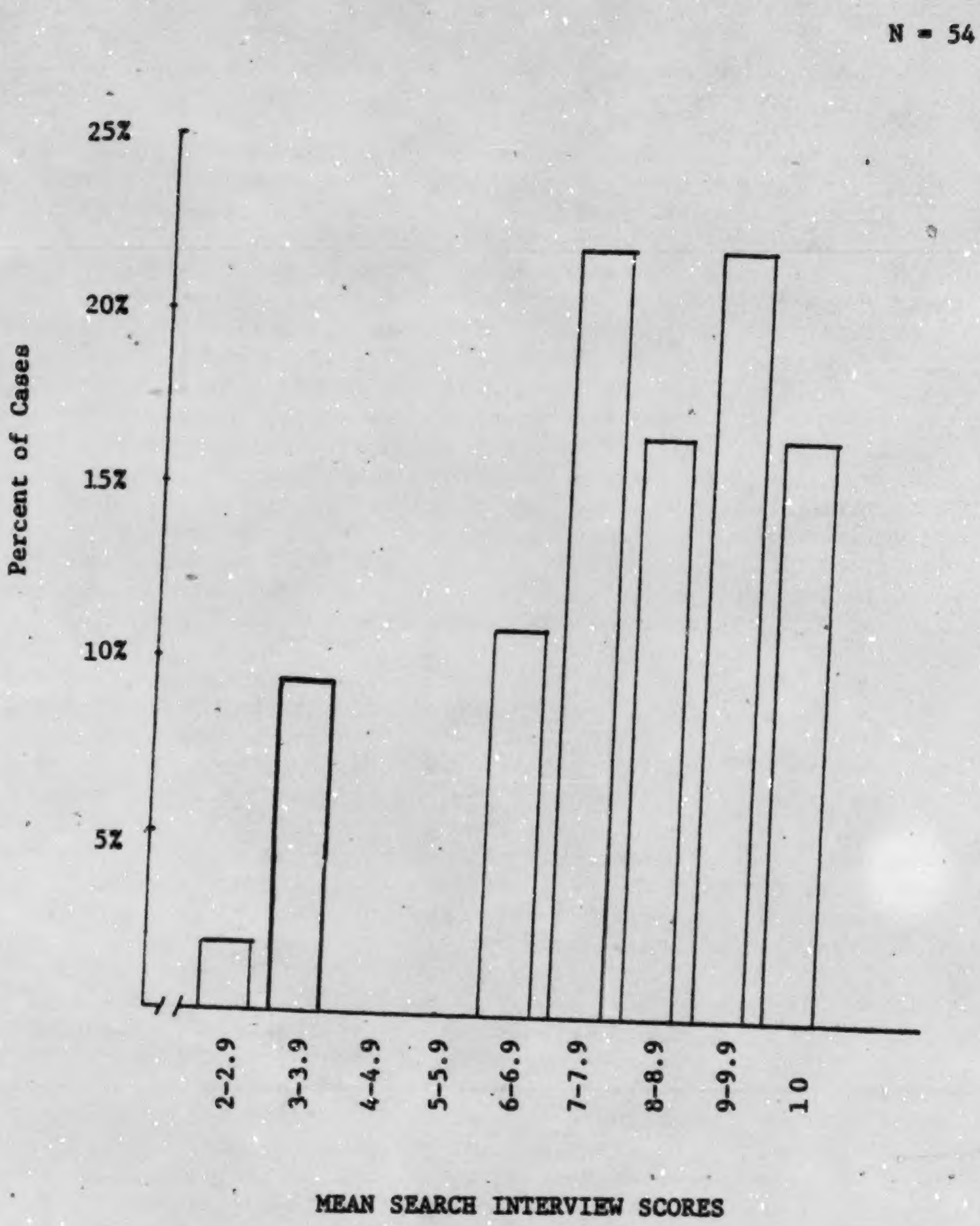


Figure 5. Distribution of Mean Scores for User Perceptions of the Search Interview.

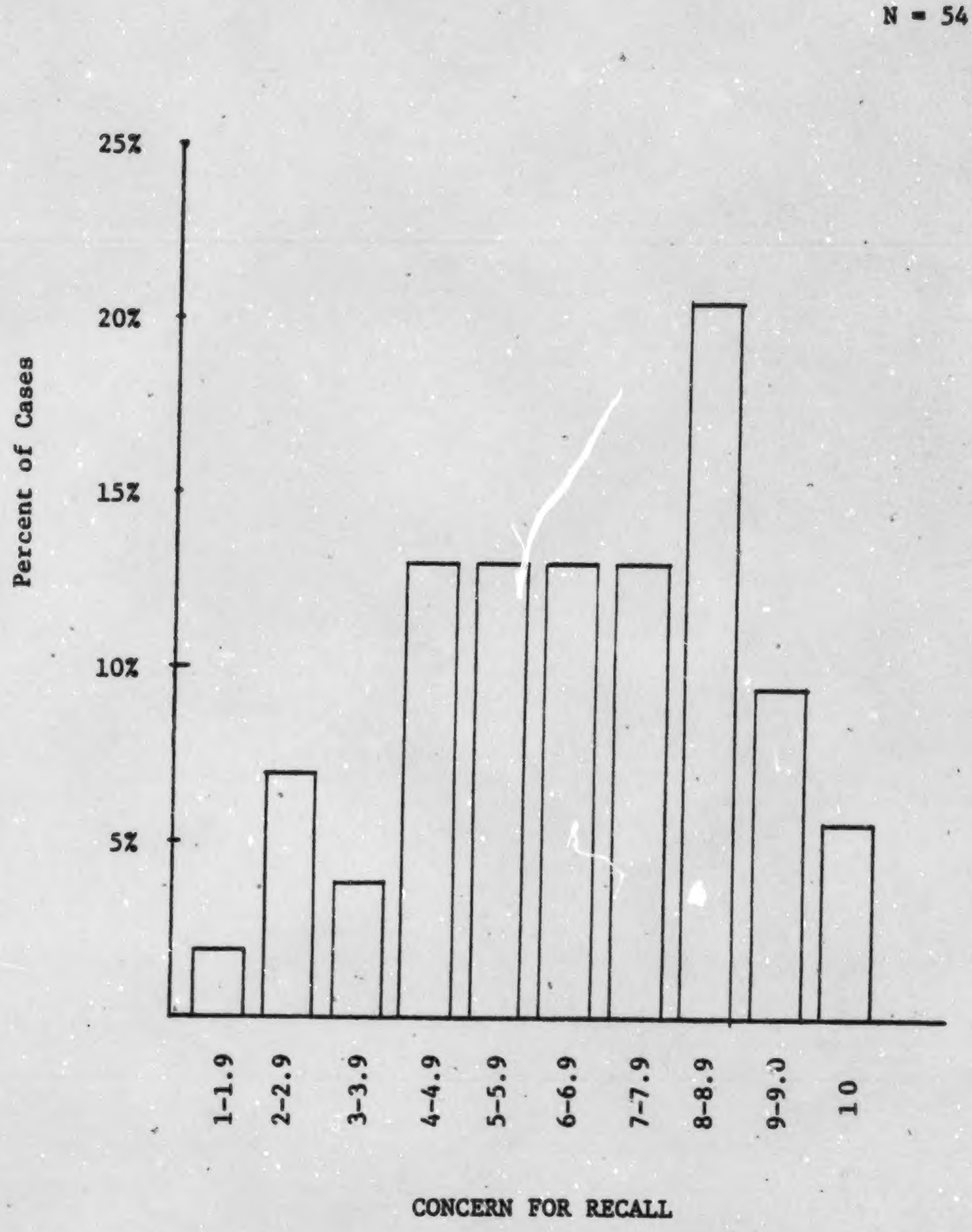
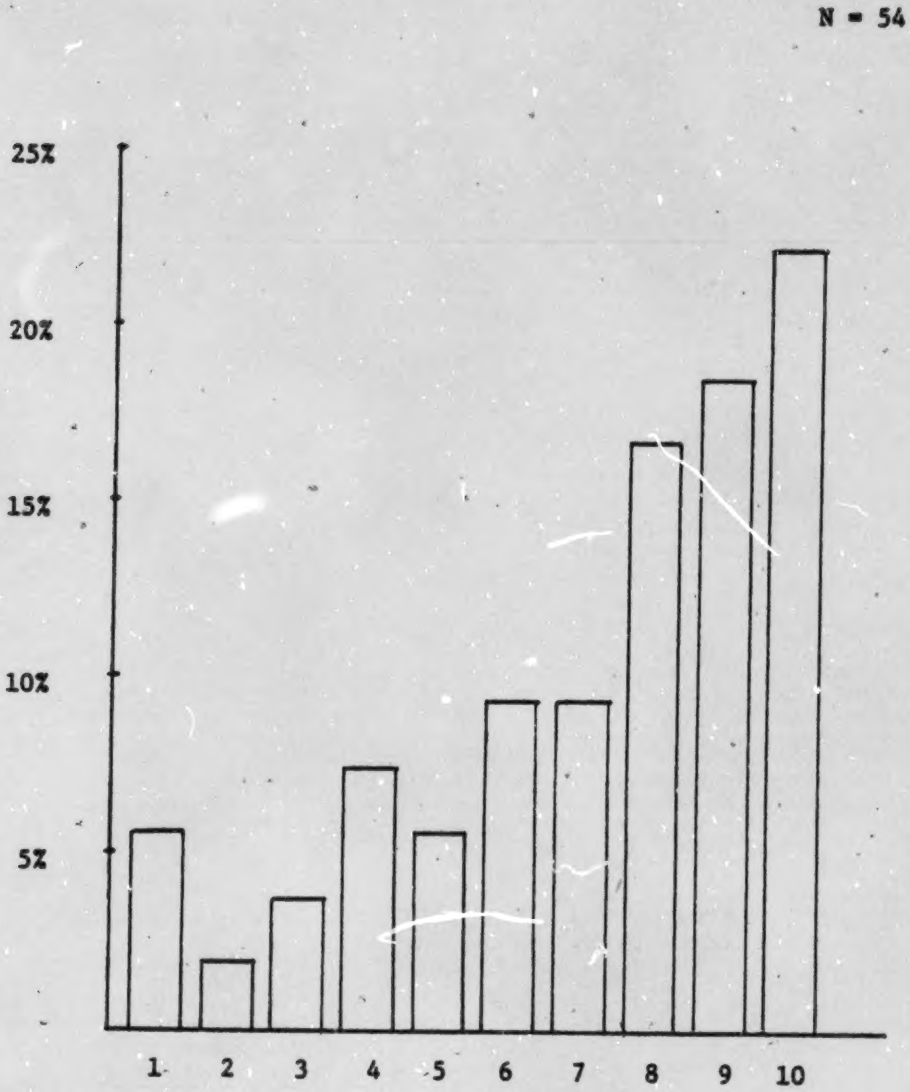


Figure 6. Distribution of Concern for Recall Scores.



SATISFACTION WITH THE PROPORTION OF RELEVANT CITATIONS

Figure 7. Distribution of Satisfaction with the Proportion Relevant Scores.

User Differences

Results were examined for differences by user type (faculty, students) in the areas of relevance, concern for recall, perception of the searcher in the interview, and information-giving activity (one-tailed tests). Differences in value and satisfaction with the proportion of relevant citations were also considered (two-tailed tests). The Mann-Whitney U procedure (Siegel, 1956, p. 116) was used in the analysis. Table 2 provides summary results.

TABLE 2

Summary Results for Testing Differences by Type of User (Faculty, Student)

Measure	Mean Rank by Group Faculty (N = 22)	 Student (N = 32)	U	P
Relevance	28.00	26.29[a]	319.0	.35
Concern for Recall	32.39	24.14	244.5	.03
Perception of Searcher	27.64	27.41	349.0	.46
User Information-Giving[b]	8.31	10.45	30.5	.20
Value	30.52	25.42	285.5	.24
Satisfaction with Proportion of Relevant Citations	31.52	24.73	263.5	.12

[a] 31 students

[b] analyzed interview — 8 faculty, 10 students

Earlier reports (Rees and Schultz, 1967: Saracevic, 1970) provided a basis for belief that, in relevance assessments, knowledgeability may be associated with more stringent assessment. In the current study these findings were extrapolated to a model which suggested that this same stringency of assessment would occur in actual searches of an online data base. Specifically, it was suggested that a more knowledgeable group of biomedical users (faculty) would indicate lower relevance scores, lower concern for recall scores (be more concerned about recall), and lower searcher perception scores than would a less knowledgeable group of users (students). Since value and satisfaction with the proportion of relevant citations retrieved had no immediately apparent connection with knowledgeability, it was suggested that the two groups would exhibit no differences in these areas. When the interview itself was considered it was thought that the more knowledgeable users would exhibit greater information-giving activity.

As anticipated, no difference in user groups was noted on the two more general measures of value and satisfaction with the proportion of relevant citations. However, in areas where it was expected that faculty users would make more stringent assessments this did not occur. Faculty and student users exhibited no significant rank order differences in the direction predicted for relevance scores, recall scores, or search interview scores. It is interesting to note that for users in this study, students actually appeared to be more concerned with recall than did faculty users (P = .03). In information-giving activity faculty users exhibited no significant rank order differences when compared with students.

Since the expectation of more stringent assessments and more information-giving activity from faculty was based on the supposition that faculty users were more knowledgeable, several items from the user request form were examined to determine if indicators of knowledgeability were present in the sample studied.

Utilizing responses from question 12 on the request form, degrees completed by the faculty and student users were examined. To the extent that knowledgeability is equated with higher levels of educational attainment, faculty users can be considered more knowledgeable when the proportion of users in each category is considered (Kendall's tau C = .94, p = $<$ .001). Eighteen students had completed one degree, the Bachelor's; eleven students had an undergraduate degree plus the Master's. These groupings were descriptive for none of the faculty. Ten users had completed a professional degree program (M.D., J.D., D.D.S.) or had two Master's degrees. Three students were in this category. Fifteen users had a Ph.D. or a Ph.D. plus another degree (e.g., M.D.). No students were included in this category.

Responses to question 13 on the request form were examined for variables related to general knowledge of, and continuing interest in, the topic of the search. No significant associaton is noted between user status and the awareness of several recent publications on the topic of the request (19 out of 21 faculty responded affirmatively, as did 23 out of 32 students), or between status and the user's continuing interest in the topic (all faculty indicated continuing interest, 30 out of 31 students indicated the same interest). However, faculty users were more likely to read on a regular basis journals on the topic of the request, and to have published in the area of the request. Table 3 indicates journal reading habits, and Table 4 indicates responses concerning publication.

Table 3

User Type by Regular Reading of Journals

	Read Journals		
User	Yes	No	
Student	13	19	32
Faculty	16	6	22
	29	25	54

$X^2 = 4.19$; $p < 0.05$

Table 4

User Type by Publication of Related Article(s)

	Have Published		
User	Yes	No	
Student	3	29	32
Faculty	8	14	22
	11	43	54

$X^2 = 4.31$; $p < 0.04$

By topic the evaluated searches could be subjectively divided into two areas, conventional biomedical searches, and those more directed to the social sciences (e.g., spirituality of patients, sexuality and the mentally retarded). No systematic relationship between type of user and area of search was apparent, and when evaluative measures were contrasted (relevance, satisfaction with the proportion of relevant citations, concern for recall, value, and perception of the searcher) there was no significant difference (p = ≤ .05) in the distribution of evaluative scores by topic area of the search. Table 5 indicates summary results for testing evaluative differences by area of search.

Table 5

Summary Results for Testing Differences

by Subject Area of Search
(Biomedical, Social Science)

	Mean Rank by Area			
Measure	Biomedicine (N = 41)	Social Science (N = 13)	U	P (two-tailed)
Relevance	27.67	24.71[a]	218.5	.56
Satisfaction with proportion relevant	29.61	20.85	180.0	.08
Concern for recall	29.57	20.96	181.5	.09
Value	28.95	22.92	207.0	.22
Perception of Searcher	27.71	26.85	258.0	.86

a = 12 cases

Faculty users were more likely to have participated in a previous online interview than were student users. Seventeen faculty users (77.2%) had participated in one or more such interviews; twelve students (37.5%) had previous experience with online interviews. However, if cases are contrasted by previous experience, or no experience there is no significant difference in the distribution of evaluative scores for the two situations. Table 6 indicates summary results for examining distribution differences by experience with online interviews.

Table 6

Summary Results for Differences by Previous Online Interview Experience (None, One or More)

	Mean Rank			
Measure	None (N = 25)	(One or More (N = 29)	U	P (two-tailed)
Relevance	25.10	28.70[a]	302.5	.40
Satisfaction with proportion relevant	26.68	28.21	342.0	.72
Concern for recall	27.64	27.38	359.0	.95
Value	27.08	27.86	352.0	.85
Perception of Searcher	29.28	25.97	318.0	.44

a = 28 cases

Previously it was mentioned that some interview sessions incorporated the terminal search as part of the interview process. There was no apparent systematic relationship between user type (faculty, student) and participation in a particular type of interview. When interviews including the terminal session are contrasted with non-terminal interviews an interesting difference in the area of the perception of the searcher in the interviews is noted. This suggests that, for this study, users in terminal session interviews assigned higher values in the area of perception of the searcher. Table 7 outlines results of examining terminal and non-terminal interviews on evaluative measures.

TABLE 7

Summary Results for Testing Differences by Type of Interview Session (Terminal, Non-Terminal)

	Mean Rank			
Measure	Terminal (N = 15)	Non-Terminal (N = 39)	U	P (two-tailed)
Relevance	33.00[a]	24.85	189.0	.09
Satisfaction with proportion relevant	32.80	25.46	213.0	.12
Concern for recall	34.07	24.97	194.0	.06
Value	29.87	26.59	257.0	.49
Perception of Searcher in Interview	37.67	23.59	140.0	.003

a = 14 cases

Information Giving, Question-Asking

This section examines the interrelationships of the user information-giving during the search interview and searcher question-asking activity with two traditional areas of evaluation concern (relevance and recall). Recall in this study is derived concept "concern for recall" rather than an actual numerical determination.

User information-giving activity scores (the sum of user activity coded for categories 4, 5, and 6 of the Bales scheme) from the 18 analyzed interviews were ranked and compared to user designations of relevance and concern for recall by means of the Kendall rank correlation coefficient. Information-giving activity was significantly related to relevance; however, the relationship to concern for recall was negligible and non-significant. Searcher question-asking activity scores (the sum of searcher activity coded for categories 7, 8, and 9 of the Bales scheme) were significantly related to user information-giving scores. Table 8 outlines the results from this examination.

TABLE 8

Correlation of User Information-Giving with
Relevance, Concern for Recall, and
Searcher Question-Asking

Measure	Kendall's tau
Relevance	.50*
Concern for Recall	.20
Searcher Question-Asking	.58**

* p = .004

** p = .001

Earlier it was suggested that information-giving by the user can be either voluntary or searcher elicited. To examine the remaining relationship between information-giving and relevance when the effects of question-asking are held constant, the Kendall partial rank correlation coefficient was determined (Siegel, 1956, p. 223). The partial rank correlation coefficient (relationship of information-giving with relevance with the effect of question-asking held constant) was .38. The decrease in the relationship initially observed (.50) provides some basis for suggesting that question-asking by the searcher moderately enhances the relationship between user information-giving scores and relevance scores.

The volume of information-giving units and question-asking units varied considerably over the evaluated interviews. The mean score for searcher question-asking units was 37.5 (SD 27.3); the mean score for information-giving units was 80.8 (SD 59.6).

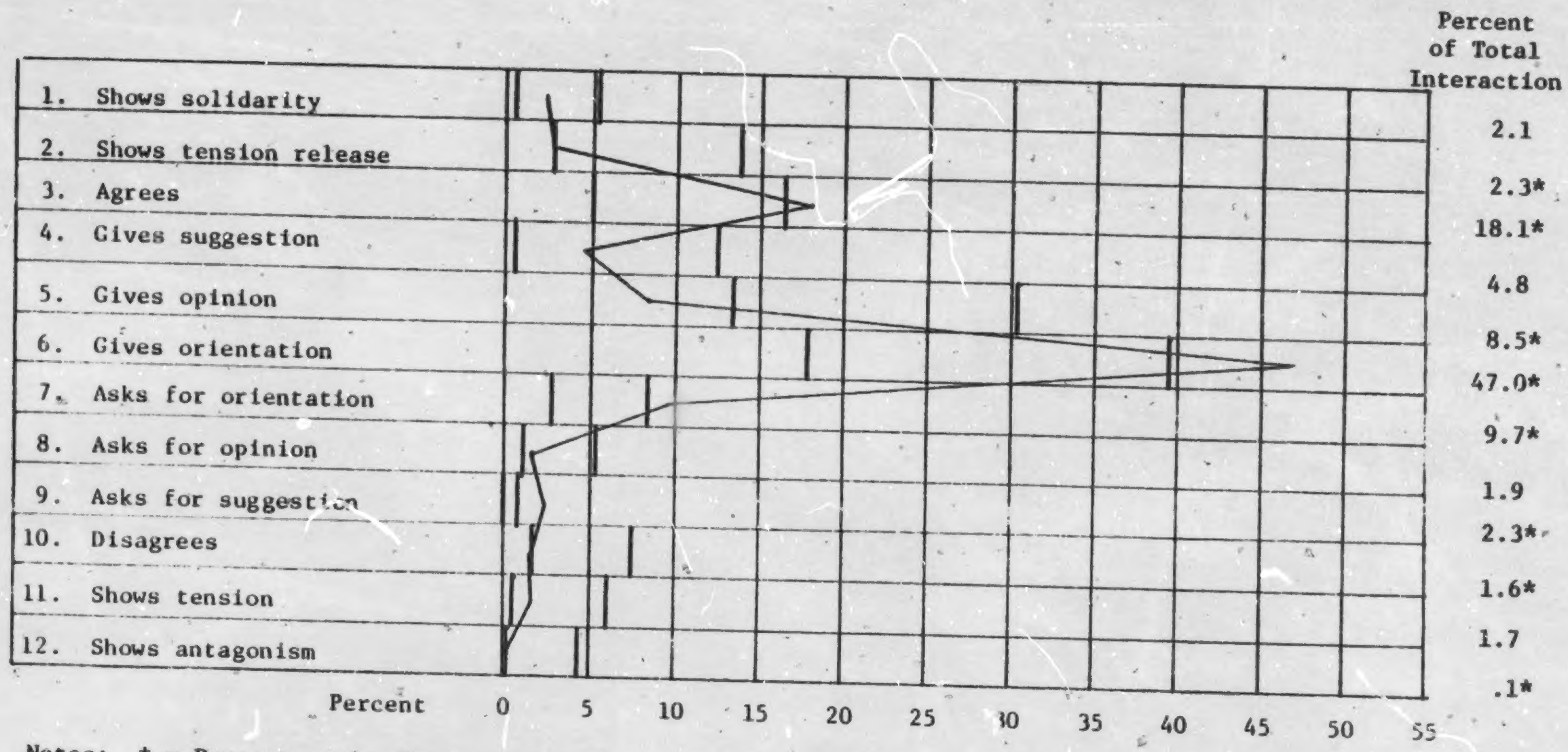
Seven of the eighteen coded interviews included terminal sessions. More question-asking was noted for these interviews ($\bar{X} = 62.7$, SD 26.3) than for non-terminal interviews ($\bar{X} = 21.5$, SD 11.1). More information-giving was noted for users in terminal searches ($X = 119.6$, SD 51.6) than in non-terminal searches ($\bar{X} = 56$, SD 52.1).

Although relevance itself has a moderate relationship with the value of the search results ($\tau = .47$) and information-giving appears to be more directly associated with relevance ($\tau = .50$) than does question-asking ($\tau = .36$), some interactive effect appears to occur, which suggests that training in question-asking may be a factor for consideration in preparation of online searchers. To the extent that question-asking is important, the impetus, or opportunity for additional questions appears more likely to arise in interactive searches at the terminal with the user present, than in interviews removed from the actual terminal search session.

Interview Profiles

In this section an overall profile for all coded searches is presented and compared with a standard summary profile which represents pooled findings from 21 studies which employed the Bales' scheme (Bales & Hare, 1965). Breakouts for searcher and user interaction, and interaction by type of interview (non-terminal, terminal) are also discussed.

Figure 8 gives the percentage of interaction by category for the MEDLINE interviews. Darker vertical grids indicate points 1 SD above or below those noted by Bales and Hare in the pooled interaction studies.



Notes: * = Percentage is above or below 1 SD of diagnostic population.
Total number of interaction units = 6558
Number of searches = 18
Vertical grid marks indicate 1 SD for the category.

Figure 8. Percent of Total Interaction by Category for 18 Transcribed MEDLINE Interviews.

The most obvious departure occurs in the information-giving category. Agreement units and asking for suggestion units are also higher (+ 1 SD) than those reported previously. This pattern would appear to be consistent with the aim of the search interview "to arrive at a clear, narrative, natural language statement of the user's information needs and to gather a number of facts and clues to be used to amplify or refine this statement" (Carmon, 1975, p.4). The profile is generally lower in those areas which indicate positive and negative emotional areas. Some of this may be attributed to the method of coding, i.e., utilization of a typed-transcript. Waxler and Mishler (1966) noted lesser use of negative categories in typed transcripts (loss of emotional tone in interchanges) but found that an overall comparison (rank order) of category use for the two methods (typed transcript, tape and typed-transcript) gave similar distributions. The lower rates noted for categories 1 and 2 may simply reflect the business-like nature of the interview situation. The higher percent of units scored for the agreement category (Category 3) parallels findings in other types of discussion groups where participants are required to reach consensus on an issue.

Figure 9 provides more illumination on who does what during the interview. It is apparent that the searcher dominates the task area of the profile (Categories 4 through 9). The information-giving activity by the searcher is most striking, since it seems that this giving of information is not necessarily the response to a high rate of question-asking activity by users. Users appear to take a relatively passive role in asking questions. If one of the aims of the interview is to reach a consensus on the needs of the user, it seems that the consensus that is reached may be somewhat influenced by the searcher. The searchers indicate higher levels of giving suggestion and opinion, while users are predominant in the agreement category.

Earlier it was noted that more information-giving and question-asking was present in terminal searches. Profiles of percent of searcher interaction by category, for total searcher interaction by interview type (Figure 10) show fairly similar proportional patterns for searcher interaction. Searchers in non-terminal interviews exhibited more proportional activity in the agreement category. Searchers in terminal interviews exhibited more proportional activity in asking for orientation, or facts from the user, but somewhat less in asking for opinions or suggestions. Terminal interviews appear to create more tension or withdrawal situations for searchers than do non-terminal searches.

Users, like searchers, also have similar interaction profiles in the two types of interview (Figure 11). Users in terminal interviews exhibit a somewhat higher proportional activity in the positive (Categories 1 through 3) and negative categories (10 through 12) than do users in non-terminal searches.

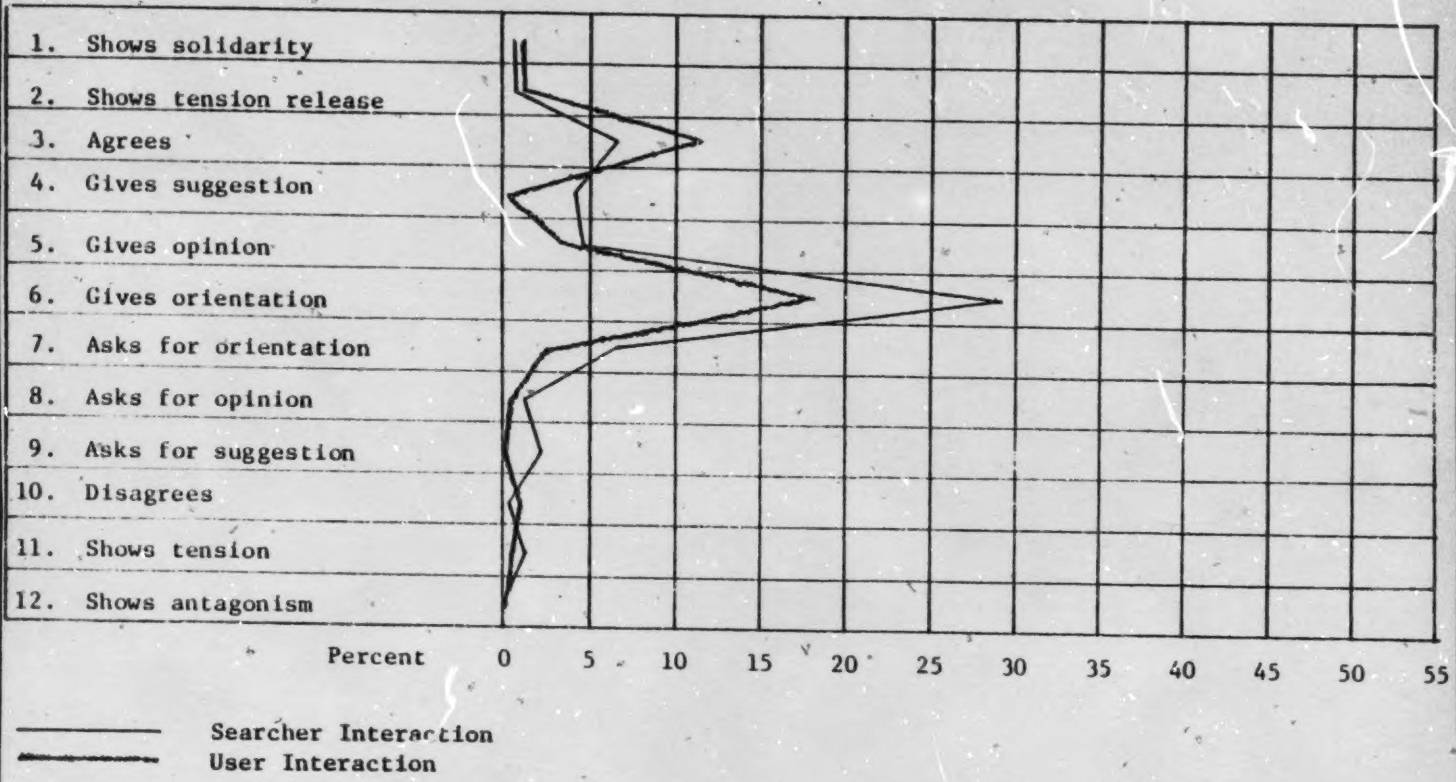


Figure 9. Searcher and User Interaction by Categories as Percent of Total Interaction for 18 Transcribed MEDLINE Interviews.

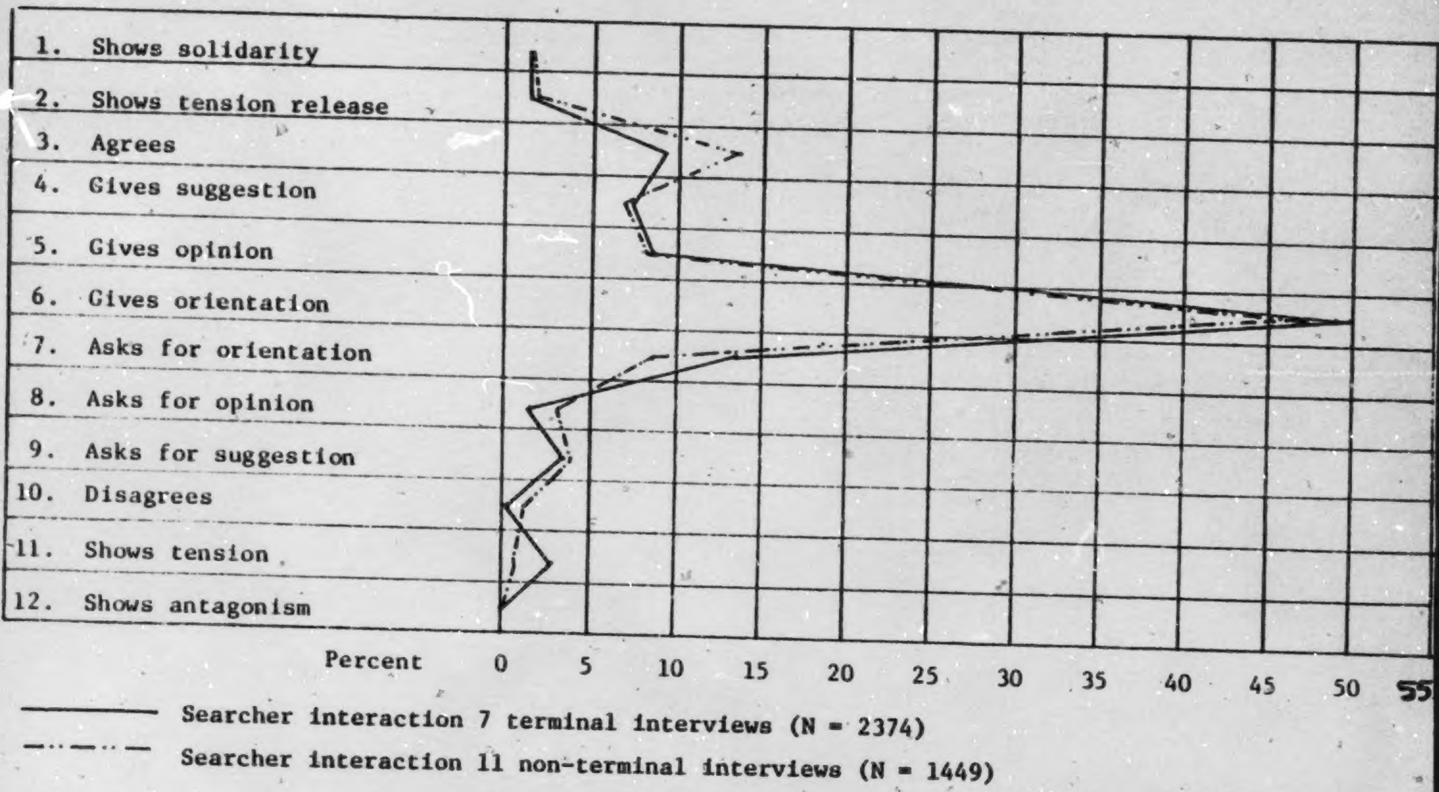


Figure 10. Searcher interaction by category as percent of total searcher interaction by type of interview.

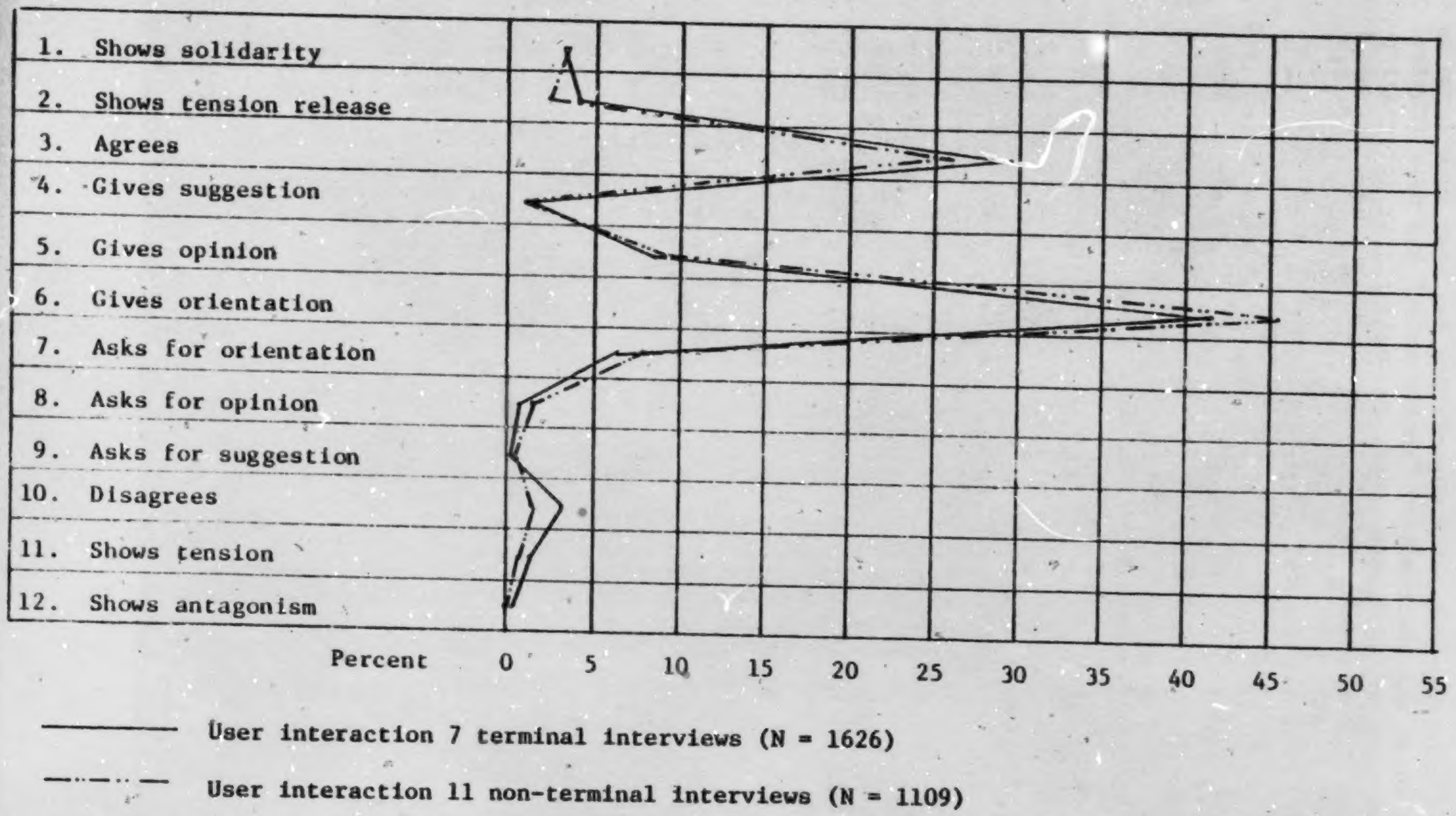


Figure 11. User interaction by category as percent of total user interaction by type of interview.

Although reference negotiation has been considered as a process in which the intermediary "interrogates" the user (Taylor, 1968), the profiles outlined here appear to suggest that information-giving is the prime activity for both searchers and users participating in an online interview. In a quantitative sense more information-giving by the user and more question-asking by the searcher appear to occur in interviews which are combined with the actual terminal search. Whether the increased activity is a function of specific events associated with terminal sessions or only the increased amount of time the interviewer and searcher spend together is not readily apparent. However, terminal sessions may provide the atmosphere for more helpful information exchanges simply because they imply an extended interview session. From a subjective perspective it seemed that users in terminal sessions volunteered several important background characteristics about themselves either while waiting for the terminal to respond to some particular input, or as a result of particular citations they viewed in the session. Searchers appeared more likely to question users about seemingly dissonant responses to a particular citation, e.g., asking why a citation was not relevant when it appeared consistent with something the user had previously asked for, or alternatively asking the user why he like a particular ciation when it did not seem to be within the scope of material previously discussed.

V. Conclusions

Conclusions resulting from this descriptive study emanate from several sources, from data tests; from trends evident in supplementary analysis of the data, and from subjective impressions developed by the investigator in the process of the study. The primary limiting factor affecting generalizability to other similar academic MEDLINE search sites, would appear to be the self-selection of users who participated in this study. Such selection is likely to be a continuing factor in the future for similar studies as very proper concerns for the privacy of individual subjects are expressed through the application of increasingly rigid procedures to protect subject rights.

For users in this study the satisfaction with the proportion of relevant citations retrieved, and expressions of concern in the area of recall for the search, appeared to be more importantly related to the value of the search results than did the actual proportion of relevant citations (the relevance score). This suggests that greater elucidation of the individual user's tolerance level for irrelevant citations at the outset of the search process may be important. This aspect is sometimes covered by a broad-narrow question on the request form (e.g., Do you want a broad search retrieving many of the relevant citations, but which may also retrieve many irrelevant citations? or, Do you want a narrow search retrieving primarily relevant citations, with few irrelevant citations, but which may exclude some relevant citations?) A question of this type would also appear to have some bearing on establishing the user's potential concern for recall. This type of question was not included on the request form for this study, so no specific examination of pre-search expressions of preference in these areas can be made. Specific elucidation of the user's preference during the interview is suggested because, in general, for interviews considered in this study, the user request form, if referred to at all, appeared to serve primarily as a source of information concerning the subject of the search. In many cases the form was not completed by the user until the end of the interview.

In a formal sense, results provide no basis for suggesting that the presumed greater knowledgeability of faculty users (with some justification for the presumption of knowledgeability evidenced in the higher degree levels attained by faculty, the greater likelihood of faculty to read on a regular basis journals related to the search topic, and the greater likelihood that they had published on the topic) is a factor leading to more stringent assessments in the areas of relevance, concern for recall, and perception of the searcher in the interview. Knowledgeability in the sense considered here also does not appear to be a factor leading to greater information-giving activity in the search interview.

Interaction profiles developed in this study showed that information-giving activity (as defined by Interaction Process Analysis) constituted the largest proportion of all searcher activity. This observation repeats in a sense the observation by Lynch that "participating librarians /in reference interviews in public libraries/ spent a significant proportion of their time giving information to patrons rather than getting information from them" (1977, p. 129). No direct examinations of searcher information giving and any assessment measures were made in this study. However, it does appear that some relationship exists between user information-giving activity and relevance scores. Question-asking by the searcher has been suggested as a contributory factor in this relationship.

Users in terminal session interviews exhibited more information-giving activity than those in non-terminal interviews. Searchers in terminal interview sessions exhibited more question-asking activity than did searchers in non-terminal sessions. A trend in the data results suggests that users participating in terminal sessions assigned somewhat higher relevance scores, satisfaction with the proportion of relevant citations scores, concern for recall scores (were less concerned about recall), and search interview scores than did users in non-terminal sessions. A subjective impression received by the investigator in listening to the tapes of the interviews is that the user's viewing of actual citations during the interview prompted interactive activity (information-giving, agreements, disagreements, and question-asking) that may have been helpful in the search process. A further examination of the type of interview (terminal, non-terminal) and assessment measures may provide additional information for improving the interactive and diagnostic process.

Bibliography

Bales, R. F. Interaction Process Analysis. Cambridge, Mass.: Addison-Wesley Press, 1950.

Bales, R. F. Interaction process analysis. In D. L. Sills (Ed.), International Encyclopedia of the Social Sciences (Vol. 7). New York: Macmillan, 1968.

Bales, R. F. & Hare, P. A. Diagnostic Use of the Interaction Profile. Journal of Social Psychology, 1965, 67, 239-258.

Benenfeld, A. R., Marcus, R. S., Pensyl, M. E., & Reintjes, J. F. NASIC at MIT, Final report, 1 March 1974 - 28 February 1975 (ESL-FR-587). Cambridge, Mass.: Massachusetts Institute of Technology, February 1975. (ERIC Document Reproduction No. ED 107 226)

Borgatta, E. F. & Crowther, B. A workbook for the study of social interaction processes. Chicago: Rand McNally, 1965.

Carmon, J. L. Model the user interface for a multi-disciplinary bibliographic information network (GN-41236). Athens, Ga.: University of Georgia, Office of Computing Activities, May 1975. (NTIS No. PB-242 964/56A)

Cohen, J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 1960, 20, 37-46.

Cooper, W. S. On selecting a method of retrieval effectiveness. Part I. The subjective philosophy of evaluation. Journal of the American Society for Information Science, 1973, 24, 87-100.

Daniels, L. A matter of form. Online, 1978, 2(4), 31-39.

Fosdick, H. An SDC-based on-line search service, a patron evaluation survey and implementation. Special Libraries, 1977, 68, 305-312.

Francillon M. Information retrieval: A view from the reference desk. Journal of Documentation, 1959, 15, 187-198.

Gothberg, H. User satisfaction with a librarian's immediate and non-immediate verbal-nonverbal communication (Doctoral dissertation, University of Denver, 1974). Dissertation Abstracts International, 1975, 35, 6839A. (University Microfilms No. 75-05025)

Gothberg, H. Immediacy: A study of communication effect on the reference process. Journal of Academic Librarianship, 1976, 2, 126-129.

Heyns, R. W. & Lippit, R. Systematic observational techniques. In G. Lindzey (Ed.), Handbook of Social Psychology (Vol. 1). Reading, Mass.: Addison-Wesley, 1954.

Hitchingham, E. E. MEDLINE use in a university without a school of medicine. *Special Libraries*, 1976, *67*, 188-194.

Holsti, O. R. *Content analysis for the social sciences and humanities*. Reading, Mass.: Addison-Wesley, 1969.

Hoover, R. E. Patron appraisal of computer-aided on line bibliographic retrieval services. *Journal of Library Automation*, 1976, *9*, 335-350.

Jahoda, G. *Instruction in negotiating the reference query*. Tallahasee, Fla.: School of Library Science, Florida State University, August 1975. (ERIC Document Reproduction Service No. ED 111 421)

Jahoda, G. *The process of answering reference questions. A test of a descriptive model*. Tallahasee, Fla.: School of Library Science, Florida State University, January 1977. (ERIC Document Reproduction Service No. ED 136 769)

Jahoda, G., Bayer, A., & Needham, W. L. A comparison of on-line bibliographic searches in one academic and one industrial organization. *RQ*, 1978, *18*, 42-49.

Kobelski, P. & Trumbore, J. Student use of online bibliographic services. *Journal of Academic Librarianship*, 1978, *4*, 14-18.

Lancaster, F. W. Evaluation of on-line searching in MEDLARS (AIM-TWX) by biomedical practicioners. *Occasional Papers of the University of Illinois Graduate School of Library Science*, 1972 (No. 101).

Librarians on MARS: session probes new machine assisted reference services. *American Libraries*, 1974, *7*, 504.

Lynch, M. J. Reference interviews in public libraries (Doctoral dissertation, Rutgers University, 1977). *Dissertation Abstracts International*, 1977, *38*, 530A. (University Microfilms No. 77-17, 547)

Lynch, M. J. Reference interviews in public libraries. *Library Quarterly*, 1978, *48*, 119-142.

McCarn, D. & Leiter, J. On-line services in medicine and beyond. *Science*, 1973, *181*, 318-324.

McCarthy, S. E., Feng, C. H., & Maccabee, S. S. Evaluation of MEDLINE service by user survey. *Bulletin of the Medical Library Association*, 1974, *62*, 367-373.

Mick, C. K. *Investigation of the public library as a linking agent to major scientific, educational, social and environmental data bases: Final evaluation report*. Stanford, Calif.: Applied Communication Research, September 1977. (ERIC Document Reproduction No. ED 148 347)

Moll, W. Observations on the AIM-TWX service at the University of Virginia Medical Library. Bulletin of the Medical Library Association, 1972, 60, 571-574.

Nunnally, J. C. Psychometric theory. New York: McGraw-Hill, 1967.

Nyren, K. The online revolution in libraries. Library Journal, February 15, 1978, pp. 439-441.

Pearse, W. H., Flora, R. E., Freeman, R., & Peeples, E. H. Medical school faculty productivity. Journal of Medical Education, 1976, 51, 201-202.

Rees, A. M. & Schultz, D. G. A field experimental approach to the study of relevance assessments in relation to document searching (2 vols.). Cleveland, Ohio: Case Western Reserve University, Center for Documentation and Communication Research, October 1967. (NTIS No. PB-176-079, PB-176-080)

Saracevic, T. The concept of "relevance" in information science: A historical review. In T. Saracevic (Ed.), Introduction to Information Science. New York: Bowker, 1970.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Somerville, A. N. The place of the reference interview in computer searching: The academic setting. Online, 1977, 1, 14-23.

Summit, R. K. & Firschein, O. Two year interim report. Investigation of the public library as a linking agent to major scientific, educational, social and environmental data bases. Palo Alto, Ca.: Lockheed, September 1976. (ERIC Document Reproduction No. ED 131 857)

Swanson, R. W. Performing evaluation studies in information science. Journal of the American Society for Information Science, 1975, 26, 140-156.

Tagliacozzo, R. The consumers of new information technology: A survey of the utilization of MEDLINE. Journal of the American Society for Information Science, 1975, 26, 294-304.

Tagliacozzo, R. Estimating the satisfaction of information users. Bulletin of the Medical Library Association, 1977, 65, 243-249.

Taylor, R. S. Question-negotiation and information seeking in libraries. College and Research Libraries, 1968, 29, 178-194-

Tessier, J., Atherton, P., & Crouch, W. New measures of user satisfaction with computer-based literature searches. Special Libraries, 1977, 68, 383-389.

Waxler, N. E. & Mishler, E. G. Scoring and reliability problems in Interaction Process Analysis: A methodological note. *Sociometry*, 1966, *29*, 38-50.

Williams, M. 1977 data-base and on-line statistics. *Bulletin of the American Society for Information Science*, 1977, *4*(2), 21-23.

Notes

1. Hitchingham, E. E. *User evaluation of online searching: Categorization of elements in user evaluation forms.* A compilation of materials prepared for a poster session at the ASIS Mid-Year Meeting, Rice University, Houston, Texas May 21-24, 1978.

2. Jestes, E. C. *CAIN online project report for the period October-December 1973.* Unpublished manuscript, Unviersity Library, University of California-Davis, June 1974.

Appendix A

CONSENT FORM

OAKLAND UNIVERSITY Rochester, Michigan 48063

KRESGE LIBRARY

A research study supported by the Office of Libraries and Learning Resources, Office of Education is being conducted at the Houston Academy of Medicine - Texas Medical Center Library, the University of Michigan and Wayne State University. This project is directed to an investigation of two areas: 1) the relationship of user characteristics and search purpose with the user judgement of MEDLINE search results and 2) an examination of the MEDLINE search interview for the identification of variables related to user judgment of search results. Primary data collected for this study include user information supplied on a MEDLINE request form; tape recorded interviews for MEDLINE searches; typed copies (all identifying characteristics deleted) of interviews from respondents completing a user evaluation form; and user evaluation forms returned by selected respondents. It is anticipated that the results of this study may provide a basis for clarification of performance objectives for on-line information retrieval systems. As a current user, and a potential future user, of such systems your cooperation in this project is requested.

Your participation involves three areas:

1. Completion of a search request form.
2. Participation in a tape-recorded interview session. Should you agree to recording, you may change your mind at any time during the interview and request that taping cease.
3. Potential selection as a respondent for a twelve-item evaluation form. Items include eleven scale responses and a count of relevant citations in your print-out.

The identities of all participants in this project will be held confidential by the investigator.

Participation in this project is voluntary, non-participation will not affect treatment of your MEDLINE request. Participants may withdraw from the project at any time.

September 6, 1978 Eileen E. Hitchingham

Eileen E. Hitchingham
Project Investigator

I AM WILLING TO PARTICIPATE IN THE PROJECT OUTLINED ABOVE.

Requester ______________ Searcher ______________
Date ______________ Date ______________

IN THE FUTURE, IF OTHER QUALIFIED INVESTIGATORS SHOULD REQUEST COPIES OF INTERVIEW TAPES AND/OR COMPLETED REQUEST FORMS, I AGREE TO THEIR RELEASE BY THE PROJECT INVESTIGATOR IF ALL NAME AND INSTITUTION REFERENCES ARE DELETED.

Requester ______________ Searcher ______________
Date ______________ Date ______________

Appendix B

REQUEST FORM

Date ____ / ____ / ____
Mo / Day / Year
Searcher ____________
Agree to Tape ______ Yes ____ No
Search Number ____________
Citations ___ On-line ____ Off-line

MEDLINE SEARCH REQUEST FORM

PLEASE PRINT

1. Name ______________________________

2. Telephone Number ______ / ______ / ______
Area Code / Number / Extension

3. Mailing Address ______________________________
____
City / State / Zip

4. Primary Institutional Affiliation ______________________________
Department ______________________________

5. Position or Title ______________________________

6. Number of previous on-line search interviews.
None ________
1 - 5 ________
6 - 10 ________
More than 10 ________

7. The results of this search are primarily for my own use.
Yes ________
No ________

8. If NO, please indicate Name, Title and Address of person for whom Search is intended:

Name ______________________________
Title ______________________________
Address ______________________________
____
City / State / Zip

9. When search is completed:

Send results by mail ______
I will pick up results ______

10. Check the ONE blank which BEST describes the requester of the search. If two designations appear to be applicable (e.g. a Physician who is also a Graduate student) check the status which relates MOST to your request for a search.

EXAMPLE:

A Physician/Graduate student initiating a MEDLINE request PRIMARILY for an actual patient care problem would designate "Physician."

A Physician/Graduate student initiating a MEDLINE request PRIMARILY for graduate-related coursework would indicate "Graduate Student."

______ 1. Undergraduate student
______ 2. Medical student
______ 3. Graduate student (Master's level)
______ 4. Graduate student (Ph.D. level)
______ 5. Instructor (includes Clinical & Adjunct appointments)
______ 6. Ass't. Professor (includes Clinical & Adjunct appointments)
______ 7. Assoc. Professor (includes Clinical & Adjunct appointments)
______ 8. Professor (includes Clinical & Adjunct appointments)
______ 9. Postdoctoral appointment
______ 10. Intern/Resident
______ 11. Academic researcher
______ 12. Physician
______ 13. Nurse
______ 14. Dentist
______ 15. Pharmacist
______ 16. Other (please specify) ______________________________

11. Check the ONE blank which BEST describes your purpose in requesting a MEDLINE search.

______ 1. Grant project (development, in progress, completion)

______ 2. Preparing an article based on research work for publication

______ 3. Preparing a review article for publication

______ 4. On-going research which will lead to publication

______ 5. Dissertation (Doctoral degree)

______ 6 Thesis (Master's degree)

______ 7. Term paper

______ 8. Class project

______ 9. Background material for a seminar

______ 10. Background material for a speech or talk

______ 11. Instruction or teaching

______ 12. Clinical application (diagnosis, treatment of a patient or client)

______ 13. Personal bibliography, no immediate application

______ 14. Other (please state) ____________________

12. Check ALL degrees completed

______ 1. High school

______ 2. Associate's degree

______ 3. Bachelor's degree

______ 4. Master's degree

______ 5. Doctoral degree

______ 6. M.D.

______ 7. D.D.S.

______ 8. Other (please state) ____________________

13. Please consider the following statements as they apply to the topic of your Search request.

1. I am aware of several recent (last 2 years) publications related to the topic of my Search request.	1.	Yes No	____ ____
2. I read on a regular basis (weekly, monthly) several journals which have articles related to the topic of my request.	2.	Yes No	____ ____
3. I have published an article or articles related to the topic of my Search request.	3.	Yes No	____ ____
4. The topic of this search request is likely to be of continuing interest to me (next one or two years).	4.	Yes No	____ ____

14. Search request: Please give a detailed statement of the subject matter for which the Search is to be conducted. Define any terms which may have special meaning relative to your Search and/or any concepts you wish to exclude.

15. **Relevant citations**: If possible list author, title and publication data for any known relevant articles published within the last three years. The citations will be used as a guide for retrieving similar citations related to your needs.

1. Author ____________________
   Title ____________________
   Publication Data ____________________
2. Author ____________________
   Title ____________________
   Publication Data ____________________
3. Author ____________________
   Title ____________________
   Publication Data ____________________

--------------------For Use in WAYNE Searches Only.--------------------

FOR LIBRARY USE ONLY. DO NOT WRITE IN THIS BOX
Part 2
Library Data

☐ WSU formulated search
Search unit(s) ______
Line time ______
Form time ______

☐ Pre-formulated search
Search unit ______
Line charges ______

☐ Institution run search
Minutes on-line ______

☐ Recurring search
☐ Monthly ☐ Quarterly

Method of payment
☐ Free ☐ Check($________)
☐ IRB ☐ Cash ($________)
☐ To be invoiced

Search formulated by ____________
Search run by ____________
Date run ____________ ☐ NLM ☐ SUNY
Taskname ____________

Total off-line prints:
MEDLINE ______@ __ per cit.
BACK72 ______@ __ per cit.
BACK69 ______@ __ per cit.
BACK66 ______@ __ per cit.
OTHER (specify) ____________________@ __ per cit.

TOXLINE ______@ __ per cit.
TOXBACK ______@ __ per cit.
CANCERLINE ______@ __ per cit.
EPILEPSY ABSTS. ______@ __ per cit.

Specify files to be searched:

☐ MEDLINE (latest 2-3 years)
☐ BACK72 (1972-74)
☐ BACK69 (1969-71)
☐ BACK66 (1966-68)

☐ TOXLINE (1971 to present)
☐ TOXBACK (1940-70)
☐ CANCERLINE (1963 to present)
☐ EPILEPSY ABSTS. (1945 to present)

☐ Other (specify) ____

____

Search limitations: Please check all boxes and supply information pertinent to the scope of your search.

1. ☐ No restrictions

2. ☐ Human only

   a. ☐ Male ☐ Female

   b. ☐ Infant, newborn (to 1 mo.)
      ☐ Infant (1-23 mos.)
      ☐ Child, preschool (2-5 yrs.)
      ☐ Child (6-12 yrs.)
      ☐ Adolescence (13-18 yrs.)
      ☐ Adult (19-44 yrs.)
      ☐ Middle age (45-64 yrs.)
      ☐ Aged (65 yrs. and older)

3. ☐ Animal only. If only certain animals or animal groups, list below:

   a. ____________________
   b. ____________________
   c. ____________________
   d. ____________________
   e. ____________________
   f. ____________________

4. Language restrictions:

   ☐ Accept all languages
   ☐ English
   ☐ English abstracts

   Other languages accepted (specify):

   ____

   ____

Search Formulation

Pertinent MeSH headings or keywords:

1. ____________________
2. ____________________
3. ____________________
4. ____________________
5. ____________________
6. ____________________
7. ____________________
8. ____________________
9. ____________________
10. ____________________
11. ____________________
12. ____________________
13. ____________________
14. ____________________
15. ____________________
16. ____________________
17. ____________________
18. ____________________
19. ____________________
20. ____________________

Search statements: FOR LIBRARY USE ONLY.

		Pstg.	Off-Line
1.			
2.			
3.			
4.			
5.			
6.			
7.			
8.			

Appendix C

EVALUATION FORM

Searcher	________
Search No.	________

MEDLINE USER RESPONSE FORM

You have been selected from a number of users who participated in a taped interview for a recent MEDLINE search. Your cooperation in completing this form as part of the follow-up analysis of user judgment of search results is sought.

The questions consider four areas:

1. Relevance -- Your judgment of each citation retrieved in the MEDLINE search and its logical connection (relationship) to the question you posed to the system.
2. Recall -- Your judgment of the "completeness" of the search, i.e., are you satisfied that the search retrieved all of the relevant citations in the MEDLINE data base?
3. Value of the Search -- How valuable was the search in meeting the information need which prompted your request?
4. Search Interview -- Your judgment concerning the ability of the searcher to assess and explore all aspects of your information need during the search interview.

Answer ALL questions as best you can and feel free to make any comments you think necessary. Your answers and comments will be kept confidential.

Please complete this form and return it within two weeks to:

Eileen E. Hitchingham
Kresge Library
Oakland University
Rochester Michigan 48063

A stamped envelope is enclosed for your use.

Please turn to questions on the following pages.

rev3

I. RELEVANT

1. Complete this question ONLY if search results (citation print-out) are no longer available to count the number of relevant citations. If the print-out is available GO TO QUESTION 2.

My estimate of the percent of relevant (related to my search question) citations retrieved in my search is:

0% ______
1 - 10% ______
11 - 20% ______
21 - 30% ______
31 - 40% ______
41 - 50% ______
51 - 60% ______
61 - 70% ______
71 - 80% ______
81 - 90% ______
91 - 100% ______

(GO TO QUESTION 3)

2. Your search retrieved ______ citations. Review each citation in your print-out and count the total number of relevant citations. Relevant citations are ALL citations which you consider to be related to your search question, thus, relevant citations include those with which you were already familiar prior to the search as well as new citations related to your question.

Number of relevant
Citations in the print-out ______

(GO TO QUESTION 3)

3. Indicate your degree of satisfaction with the proportion of relevant citations retrieved in the search by marking an X on the appropriate space below.

Unsatisfactory ___:___:___:___:___:___:___:___:___:___ Satisfactory
1 2 3 4 5 6 7 8 9 10

rev3

(COMPLETE ALL FOLLOWING QUESTIONS)

II. RECALL

Indicate your level of agreement with the following statements by marking an X in the appropriate space.

1. I believe that the search retrieved most of the relevant citations in the MEDLINE data base.

   Disagree ___:___:___:___:___:___:___:___:___:___ Agree
   1 2 3 4 5 6 7 8 9 10

2. In assessing my search results I am concerned because there is no way to judge completeness.

   Disagree ___:___:___:___:___:___:___:___:___:___ Agree
   1 2 3 4 5 6 7 8 9 10

3. I am concerned because the search results omitted relevant citations with which I was familiar prior to the search.

   Disagree ___:___:___:___:___:___:___:___:___:___ Agree
   1 2 3 4 5 6 7 8 9 10

4. The search results included fewer relevant citations than I expected.

   Disagree ___:___:___:___:___:___:___:___:___:___ Agree
   1 2 3 4 5 6 7 8 9 10

III. SEARCH VALUE

1. Consider the purpose for which you requested this search. Indicate your value assessment of the search results (the ability of the search to meet the need prompting your request) by marking an X in the appropriate space.

   No Value ___:___:___:___:___:___:___:___:___:___ Major Value
   1 2 3 4 5 6 7 8 9 10

IV. SEARCH INTERVIEW

Indicate your level of agreement with the following statements by marking an X in the appropriate space.

1. The searcher was knowledgeable concerning the use of the MEDLINE data base for my question.

   Disagree ___:___:___:___:___:___:___:___:___:___ Agree
   1 2 3 4 5 6 7 8 9 10

2. I feel that the searcher understood my request after the interview.

   Disagree ___:___:___:___:___:___:___:___:___:___ Agree
   1 2 3 4 5 6 7 8 9 10

3. The searcher understood my purpose in initiating the request.

   Disagree ___:___:___:___:___:___:___:___:___:___ Agree
   1 2 3 4 5 6 7 8 9 10

4. The searcher was thorough in exploring all aspects of my search question.

   Disagree ___:___:___:___:___:___:___:___:___:___ Agree
   1 2 3 4 5 6 7 8 9 10

5. The searcher suggested terms appropriate to the subject of my request.

   Disagree ___:___:___:___:___:___:___:___:___:___ Agree
   1 2 3 4 5 6 7 8 9 10

V. COMMENTS

THANK YOU FOR YOUR COOPERATION IN COMPLETING THIS FORM.

Do you wish to receive a report on the final project results?

Yes ___
No ___

rev3

Appendix D

PROCEDURES

MEDLINE PROJECT PROCEDURES

TEXAS MEDICAL CENTER LIBRARY

Supplies

Materials needed for the project include:

1. Blank cassettes
2. Tally sheets for MEDLINE Searches not included in the project
3. Consent/Request Forms
4. Tape recorder (available at Texas Medical Center Library)
5. Sample Evaluaton Forms
6. Return envelopes for cassettes and Consent/Request forms

Population to be Recorded

For the purpose of this project it is desirable to seek participation from all users who come to the Library and request a MEDLINE Search.

Tally Sheets for Non-Recorded MEDLINE Searches

Since I would like to know the proportion of recorded MEDLINE Searches to all MEDLINE Searches conducted during the participation period, tally sheets are available to note each MEDLINE Search completed for which there is no recording. On the tally sheet please note reason, date, and searcher for any unrecorded MEDLINE Search you may do.

Sequence of Events

1. Consent

When an individual requests a MEDLINE Search, in person, explain that the Library is participating in a MEDLINE evaluation study and you are inviting them to participate in the project. Give the requester a Consent/Request form for her/his consideration. If the requester agrees, obtain her/his signature(s) on the form. (Some requesters may prefer to sign only the first release, others may sign both statements).

Sequence of Events - continued

2. Recording

   Begin recording on Side #1 of the cassette immediately after the consent Form is signed. Indicate:

   - Date
   - Requester Name
   - Time
   - Search Number (at top of Consent Form)
   - Searcher Name

3. Request Form

   a) Have user fill out Request Form down to the dotted line on page 5. Additional pages (6-7) were added for the convenience of Wayne State University; they can be cut off if they are not useful for you).

   b) Check to see that all information is completed, particularly name, mailing address and phone number.

4. Interview

   a) Conduct interview according to usual method

   b) If interview is longer than 30 minutes, turn cassette over and continue recording

   c) When interview is completed indicate time and stop recording

   d) If interview includes working with the requester at the terminal this interaction should be recorded also.

5. After Interview

   a) Label cassette with

   - date
   - search number
   - requester name
   - searcher name or initials

   b) Check to see that Consent Form has requester signature

   c) Give the requester a SAMPLE evaluation form. Indicate that she/he may be contacted in the future to respond to a similar questionnaire. The user is NOT to return the sample, it is for her/his information only.

Sequence of Events - continued

5. After Inverview - continued

   d) Sign Consent Form in searcher signature position. If, for some reason, you do not wish to sign even though the requestor has agreed to participate, it would be helpful if you would note this on the signature form so that I will know that it has not just been forgotten.

6. After Search

   a) If on-line citations are printed, note the number of citations in the box in the upper right corner of the request form.

   b) If off-line citations are sent to the user, note the total number of off-line citations sent to the user (e.g. MEDLINE and nay backfiles).

   c) If both on-line and off-line citations result from the search note both of these numbers in the appropriate spot in the box in the upper right of the request form.

   d) Once a week mail all completed cases (the tape plus the Consent/Request form which notes the number of citations given to the user) in the pre-addressed envelopes.

Appendix E

COVER LETTER

OAKLAND UNIVERSITY Rochester, Michigan 48063

KRESGE LIBRARY

Dear

As you recall, you recently participated in a taped interview for a MEDLINE Search Request. With an Office of Education Grant I am conducting a study of user evaluation of on-line searches; you have been selected as a MEDLINE user important to this study. I am requesting your support in completing the enclosed questionnaire. Since it is anticipated that this study will clarify some factors involved in user satisfaction with an on-line system, I hope you will contribute a few minutes to this project. The confidentiality of your reply will be maintained.

If you have any questions concerning the form, please call me at the number below. I appreciate your time and cooperation, and am looking forward to receiving your completed questionnaire.

Sincerely,

Eileen E. Hitchingham
On-Line Project Director

313/375-0343

k
enc

Appendix F

CODING RELIABILITY

Coding Reliability

Several earlier studies which have involved coding of events occurring in the reference or bibliographic search interview have either made no specific coder reliability determinations (Carmon, 1975), or have reported reliability in terms of simple percent of agreement by coders (Gothberg, 1975; Lynch, 1977). Holsti (1969) notes that this method, although often used, has been criticized because it fails to account for the extent of coder agreement arising from chance. For this reason a method noted by Waxler and Mishler (1966) was utilized in the present study. Cohen's "k" is a coefficient of agreement for nominal scales which expresses the proportion of agreement after chance agreement has been removed from consideration (Cohen, 1960).

The formula for computing k is:

$$k = \frac{fo - fc}{N - fc}$$

where k = the coefficient of agreement
fo = the sum of agreements observed for each category
fc = the sum of agreements expected by chance for each category
and N = the total number of observations

An approximation to the standard error of k is given by:

$$\sigma_k = \frac{\sqrt{fo\ (1 - fo/N)}}{N - fc}$$

With a large N ($\geq$ 100) the sampling distribution of k approximates normality, so that confidence limits can be expressed as:

$$95\%\ \text{confidence limits} = k \pm 1.96\ \sigma_k$$

In the present study (N = 129) a k of .77 is noted (95% confidence interval = .67 ⟷ .87). If the more commonly used percent agreement method were applied to the observations, the percent of agreement would equal the upper limit of the interval (.87).

Appendix G

USER REQUESTS: ANALYZED INTERVIEWS

Search Request Statements Presented by Users with Analyzed Search Interviews

I am interested in recent literature on reproduction in nematodes (an animal phyla) and nematomorphs. In particular the areas of spermatogenesis, oogenesis, and accessory glands in these reproductive processes. Most of these studies should be at the ultrastructural level (i.e., electron microscopy).

Cardiac rehabilitation--(exercise training programs) relative to the following diagnoses: myocardial infarction, angina pectoris, coronary insufficiency. This is to include all articles concerning rehabilitative programs which have been established. It is not to be limited by disciplines involved. All foreign language articles are to be excluded.

Blood flow measurement in 1. ovary 2. uterus 3. testis 4. effect of interuterine devices, and

The effect of educational programs on incidence of VD, heart disease and breast cancer. (E.g., is the reported incidence higher or lower--or same--after patients or public's exposure to ed. of the diseases or states?)

I would like to find articles about chlorocruorins for use in my oral examination and for my dissertation.

Clinical pharmacology of propranolol (uses and indications). Back to 1974.

Project summary: To check current literature for enlightening information, both historical and current, relevant to the areas of exercise electrocardiography. This especially entails any changes in the electrocardiogram, either at rest or during stress, both immediately prior to and one month following myocardial infarction. Of particular interest are changes in EKG after the subject has undergone an exercise program.

Effect of increased plasma K^+ on release of catecholamines from adrenal medulla.

The transport of calcium in muscle tissue as it relates to neuromuscular disease.

The subject matter is on the relationship between apgar scores and length of labor.

Milk composition x age.

Respiratory rate vs. respiratory failure.

*Transport mechanisms and transport of newly synthesized proteins into mitochondria.

---

*From tape; user did not write out a statement.

Partial denture design. Effect of occlusal-functional forces on P. dentures. Use of precision attachment in P.D. Effect on abutments in P.D.

*Epithelial neoplasms.

Dihydrofolate reductase: the enzyme, its chemistry, isolation, etc., and genetics. Also inhibitors, in research, not medical uses.

Health behavior of patients--patient compliance, patient attitudes toward health, patient health beliefs.

Beta adrenergic blocking drugs/exercise. Propanolol. pindolol. Metoprolol. Aebutolol.

---

*From tape; user did not write out a statement.

Appendix H

CODED INTERVIEW

S-1

S: OK, /6/ if you would tell me what you'd like. /8/

U: I want to get a search on beta adrenergic blocking drugs. /6/

S: Uh huh. /3/

U: ...particularly with reference to exercise /6/ and exercise response from the use of the beta adrenergic blocking drugs. /6/

S: OK, /3/ do you want to list some of those drugs for me? /8/ I don't know if they all get lumped together. /5/

U: Um, well particularly propranalol. /6/

S: P-R-O-P-? /7/

U: R-A-N /6/ Propranalol. /6/ Pindolol /6/ P-I-N-D-O-L-O-L /6/

S: Um hmm /3/

U: Metoprolol /6/ M-E-T-O-P-R-O-L-O-L /6/

S: Let me see if those are listed together. /4/

U: They should be listed under beta adrenergic /6/ or beta adrenoreceptor blocking drugs /6/ or maybe under adrenergic blocking drugs. /6/

S: Yeah /3/ OK, /6/ do you want adrenergic beta reception blockaders? /7/

U: Yes /3/ that's it. /6/

S: OK, /6/ so you want these. /9/ Let me show you what's under there /4/ so then you can decide whether you want them all of 'em or some of them. /4/

U: Um hm. /3/

S: OK, /6/ there are those that get listed. /6/

U: Um hmm. /3/

S And you can have all of them /6/ or some of them. /6/

S-2

U: Um hmm. /3/

S: It goes on /6/

U: Yeah /3/ Um ..... /11/ I probably may as well get, um all of them, /5/ because a lot of papers have, like four of them together. /6/

S: Um hm. /3/ OK. /3/

U: So it's hard to tell from a title which ones are left out, /6/ there are four in particular that I want to look at, /6/ but we may as well get them all. /6/ I particularly want to look at the work that's been done in examining exercise and the use of these drugs /6/ because there will be many more publications /6/ than there will be publications with exercise. /6/

S: Right. /3/ Right. /3/ There's usually a lot written on ... /5/ Now /6/ you're just interested in somebody who has ingested this drug and the effect of exercise on its metabolism? /7/ or what? /8/

U: Well the effects of exercise upon blood pressure /6/ and pulse, /6/ exercise tolerance, /6/ anything like that. /6/

S: OK. /3/

U: But they won't have that in the title. /6/

S: Well if I just take any, any, um, incidence of the word exercise appearing with any of these? /7/

U: Yeah. /3/ That should, /5/ that's right, /6/ the title should have, very often, in the majority of incidences, should contain the word exercise. /5/

S: Um hmm /3/ What is the response to exercise? /7/ Is it metabolized differently? /7/ or? /8/

U: Well no, /10/ it affects maximum pulse rate /6/ and the blood pressure response. /6/

S: OK. /3/ How far back do you want to go back on this? /8/

U: Um, well, you have different categories don't you? /8/

S: Right. /3/ The current file is January 76 forward /6/ but the file itself can go back as far as 1966. /6/

S-3

U: Um, what's the differential in price between the ...? /7/

S: OK, /6/ the current file is either a $5.00 /6/ or a $7.00 search /6/ depending on whether or not you want abstracts /6/ if you go all the way back it's $18.00, /6/ or $22.00 if you want abstracts, /6/ and then in between it ranges between those two /6/ so if you want to go back to 1972 /6/ it's $14.00 /6/ or $18.00 depending on the abstracts. /6/

U: Let's go back to '72 /4/ I have a paper with my account number on it. /6/

S: OK, /3/ OK. /3/ And do you want um human and animal both? /9/ Are you going to be using...? /8/

U: Uh, just human. /6/

S: Do you want just English? /7/ or all the languages? /9/

U: Just English. /6/

S: OK. /3/ Do you have a sense of how much stuff is being written on this? /8/

U: Um..... /11/

S: Like 50 articles a year? /9/ or more? /9/ or less? /9/

U: I imagine it would be considerable, /5/ it probably would be 50 articles a year at least. /5/

S: OK. /3/

U: Maybe more than that. /5/ Some of the drugs have been studied much more intensively than others. /6/

S: Um hmm. /3/ OK. /3/ OK, /6/ and your affiliation is with ...? /7/

U: (gives affiliation) /6/

S: Department? /7/

U: Medicine. /6/

S: OK. /6/ And you have an account number there? /7/

S-4

U: This. /6/

S: OK, /6/ what's a phone number I can reach you at? /7/

U: (gives number) /6/ Now I'm going to be away later in the week, /6/ but my secretary can come and get it. /6/

S: OK, /6/ This takes about a week or 10 days /6/ unless you want the first 10 like tomorrow afternoon. /6/

U: That's OK, /3/ I'm going to be away /6/ so next week is fine. /1/

S: OK. /3/ Did you want abstracts with it? /7/

U: No /10/ No. /10/

S: OK. /3/ OK. /3/

U: That would make it too numerous like /6/

S: My name is _______________ if you need to get back to me in the next day or so. /6/

U: OK. /3/ Right. /3/

S: If I have a problem you'll be around tomorrow? /7/

U: Yes. /3/ till Thursday. /6/

S: OK, /6/ I'll just get back to you before then. /6/

U: OK. /3/

S: If it doesn't seem quite... /5/ OK, /6/ and thank you for ... /1/

U: OK, /3/ Thanks. /1/